

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 October 2000 (12.10.2000)

PCT

(10) International Publication Number
WO 00/60359 A3

(51) International Patent Classification⁷: **G01N 33/68, C07K 7/08, 14/47, A01K 67/027, C07K 16/18, A61K 38/10, C12N 9/12**

(21) International Application Number: **PCT/GB00/01245**

(22) International Filing Date: **31 March 2000 (31.03.2000)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
9907687.9 1 April 1999 (01.04.1999) GB

(71) Applicant (*for all designated States except US*): **KUDOS PHARMACEUTICALS LIMITED [GB/GB]; 327 Cambridge Science Park, Milton Road, Cambridge CB4 4WG (GB).**

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **JACKSON, Stephen, Philip [GB/GB]; 45 Thornton Road, Girton, Cambridge CB3 0NP (GB). GELL, David, Anthony [GB/GB]; 15 Treedsdale Road, Harrogate, North Yorkshire HG2 0LX (GB).**

(74) Agents: **WALTON, Sean, M. et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).**

(81) Designated States (*national*): **AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.**

(84) Designated States (*regional*): **ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).**

Published:

— *With international search report.*

(88) Date of publication of the international search report:
28 June 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/60359 A3

(54) Title: **INTERACTIONS OF KU POLYPEPTIDES**

(57) Abstract: Ku is a protein found in a wide range of organisms. It comprises two tightly-associated subunits termed Ku70 and Ku80. The present invention relates to the discovery and characterisation of an interaction between Ku70 and Ku80 and DNA-PK_{cs}. Various applications based on this interaction are provided. These are relevant to numerous cellular processes which are of interest in therapeutic contexts.

BEST AVAILABLE COPY

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 00/01245

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/68 C07K7/08
A61K38/10 C12N9/12

C07K14/47 A01K67/027 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, BIOSIS, CHEM ABS Data, STRAND

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category ^o	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 30903 A (CANCER RES CAMPAIGN TECH ;DOWNS JESSICA ANNE (GB); JACKSON STEPHEN) 16 July 1998 (1998-07-16) claims 11-14 ---	1-12, 15-41
X	WU X ET AL: "PROTEIN-PROTEIN AND PROTEIN-DNA INTERACTION REGIONS WITHIN THE DNA END-BINDING PROTEIN KU70-KU86" MOLECULAR AND CELLULAR BIOLOGY, US, AMERICAN SOCIETY FOR MICROBIOLOGY, WASHINGTON, vol. 16, no. 9, September 1996 (1996-09), pages 5186-5193-93, XP000877319 ISSN: 0270-7306 abstract --- -/-	1-12, 15-41

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

20 September 2000

Date of mailing of the international search report

11/10/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Hart-Davis, J

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 00/01245

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOIKE MANABU; MIYASAKA TAKASHI; MIMORI TSUNEYOSHI; SHIOMI TADAHIRO: "Subcellular localization and protein-protein interaction regions of Ku proteins" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 252, 27 November 1998 (1998-11-27), pages 679-685, XP002143822 abstract --- CARY ROBERT B; CHEN FANQING; SHEN ZHIYUAN; CHEN DAVID J: "A central region of Ku80 mediates interaction with Ku70 in vivo" NUCLEIC ACIDS RESEARCH, vol. 26, 15 February 1998 (1998-02-15), pages 974-979, XP002143823 abstract --- GELL DAVID; JACKSON STEPHEN P: "Mapping of protein-protein interactions within the DNA-dependent protein kinase complex" NUCLEIC ACIDS RESEARCH, vol. 27, 1 September 1999 (1999-09-01), pages 3494-3502, XP002143824 the whole document --- SINGLETON B K; TORRES-ARZAYUS M I; ROTTINGHAUS S T; TACCIOLI G E; JEGGO P A: "The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit" MOLECULAR AND CELLULAR BIOLOGY, vol. 19, May 1999 (1999-05), pages 3267-3277, XP000929550 the whole document --- WEDRYCHOWSKI, ANDRZEJ ET AL: "Identification of proteins binding to interferon-inducible transcriptional enhancers in hematopoietic cells" J. BIOL. CHEM. (1992), 267(7), 4533-40 , XP002143825 page 4535, column 2, paragraph 2; figure 3 --- MYUNG, KYUNGJAE ET AL: "KARP-1: a novel leucine zipper protein expressed from the Ku86 autoantigen locus is implicated in the control of DNA-dependent protein kinase activity" EMBO J. (1997), 16(11), 3172-3184 , XP002143826 page 3176; figure 5 ---	1-12, 15-41 1-12, 15-41 1-41 1-12, 15-41 13-24, 29-34, 36-39 13-24, 29-34, 36-39
		-/-

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 00/01245

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MIMORI, TSUNEYO ET AL: "Isolation and characterization of cDNA encoding the 80-kDa subunit protein of the human autoantigen Ku (p70/p80) recognized by autoantibodies from patients with scleroderma-polymyositis overlap syndrome" PROC. NATL. ACAD. SCI. U. S. A. (1990), 87(5), 1777-81 , XP002143827 page 1779; figure 3 ----	13-24, 29-34, 36-39
X	WASSERMAN ROBERT; GALILI NAOMI; ITO YOSHINORI; REICHARD BETTY ANNE; SHANE SARA; ROVERA GIOVANNI: "Predominance of fetal type DJ-H joining in young children with B precursor lymphoblastic leukemia as evidence for an in utero transforming event" JOURNAL OF EXPERIMENTAL MEDICINE, vol. 176, 1992, pages 1577-1581, XP000933767 the whole document ----	13-24, 29-34, 36-39
X	US 5 635 597 A (BARRETT RONALD W ET AL) 3 June 1997 (1997-06-03) table 1 ----	13-24, 29-34, 36-39
X	DATABASE EMEST_ROD3 EBI, Hinxton, UK; Accession Number AI181344, 13 October 1998 (1998-10-13) MARRA M ET AL.: "Mus musculus cDNA clone IMAGE:1395081 5' similar to TR:Q62084 Q62084 PHOSPHOLIPASE C NEIGHBORING " XP002143829 the whole document -----	13-24, 29-34, 36-39

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 23,31,34,36,38,40-41 (in part)

Present claims 23,31,34,36,38, as well as 40 and 41 in so much as they depend on claim 23, relate to a compound defined by reference to a desirable property, namely its ability to bind to a peptide derived from polypeptides of the Ku70 or Ku80 families, binding being determined by an *in vitro* (claims 1-9) or modelling (claim 35) screening method.

No technical features of the compounds are present in the above-mentioned claims which would lead to this desirable property, the technical features formulated so as to permit the execution of a meaningful search. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found for only a very small proportion of the substances which could fall within the scope of these claims. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. No means are present in the above-mentioned claims by which compounds known in the prior art could be distinguished from novel substances. No definition of the subject matter for which protection is sought is therefore derivable from these claims (Article 6 PCT) or the description (Article 5 PCT). Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search for claims 23,31,34,36,38, 40 and 41 has been restricted to the substances which are clearly defined and supported by the description, namely polypeptides of the Ku70 or Ku80 families and peptides derived therefrom such as EGGDVDDLLDMI.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 00/01245

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9830903	A 16-07-1998	AU	5568198 A	03-08-1998
		AU	5568498 A	03-08-1998
		EP	0966683 A	29-12-1999
		EP	0965040 A	22-12-1999
		WO	9830902 A	16-07-1998
		GB	2321702 A,B	05-08-1998
		GB	2322193 A,B	19-08-1998
		GB	2329469 A,B	24-03-1999
		GB	2329248 A,B	17-03-1999
US 5635597	A 03-06-1997	NONE		

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
12 October 2000 (12.10.2000)

PCT

(10) International Publication Number
WO 00/60359 A3(51) International Patent Classification⁷: G01N 33/68. C07K 7/08, 14/47, A01K 67/027, C07K 16/18, A61K 38/10, C12N 9/12

(21) International Application Number: PCT/GB00/01245

(22) International Filing Date: 31 March 2000 (31.03.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9907687.9 1 April 1999 (01.04.1999) GB

(71) Applicant (for all designated States except US): KUDOS PHARMACEUTICALS LIMITED [GB/GB]; 327 Cambridge Science Park, Milton Road, Cambridge CB4 4WG (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): JACKSON, Stephen, Philip [GB/GB]; 45 Thornton Road, Girton, Cambridge CB3 0NP (GB). GELL, David, Anthony [GB/GB]; 15 Treesdale Road, Harrogate, North Yorkshire HG2 0LX (GB).

(74) Agents: WALTON, Sean, M. et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
28 June 2001(48) Date of publication of this corrected version:
8 November 2001(15) Information about Correction:
see PCT Gazette No. 45/2001 of 8 November 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/60359 A3

(54) Title: INTERACTIONS OF KU POLYPEPTIDES

(57) Abstract: Ku is a protein found in a wide range of organisms. It comprises two tightly-associated subunits termed Ku70 and Ku80. The present invention relates to the discovery and characterisation of an interaction between Ku70 and Ku80 and DNA-PK_{cs}. Various applications based on this interaction are provided. These are relevant to numerous cellular processes which are of interest in therapeutic contexts.

INTERACTIONS OF KU POLYPEPTIDES

The present invention relates to assays, screening methods, peptides, mimetics, and methods of use based on the discovery 5 and characterisation of an interaction between Ku70 and Ku80 and DNA-PK_{cs}. More particularly, aspects of the invention are based around peptide fragments of Ku70 and Ku80. The invention relates to numerous cellular processes which are of interest in therapeutic contexts.

10

Ku is a protein that is found in a wide range of organisms, ranging from *Saccharomyces cerevisiae* to man (Dynan and Yoo, 1998). It is expressed in all human tissues examined. Ku comprises two tightly-associated subunits of about 69 kDa and 15 about 83 kDa. These are termed Ku70 and Ku80 (or Ku86), respectively. Although some information has been obtained regarding the regions of the Ku polypeptides that interact with one another (Cary et al., 1998; Jin and Weaver, 1997; Koike et al., 1998; Osipovich et al., 1997; Wang et al., 20 1998a; Wang et al., 1998b; Wu and Lieber, 1996), little is known about the precise sites of interaction and the molecular mechanism underlying it.

The most highly characterised function of Ku at the 25 biochemical level is its ability to bind avidly to certain disruptions of the DNA double helix in a sequence independent fashion. The most well studied example of such a disruption is the DNA double-strand break (DSB; Blier et al., 1993;

Devries *et al.*, 1989; Mimori and Hardin, 1986). Other discontinuities that are recognised by Ku include single-strand breaks in the sugar-phosphate backbone of double-stranded DNA (dsDNA), and DNA single-strand to double-strand transitions, such as those that occur in hairpin loops or single-stranded gaps in a dsDNA molecule (Blier *et al.*, 1993; Falzon *et al.*, 1993). Once bound to a dsDNA end, Ku can move to internal positions in the DNA in an ATP-independent fashion (Paillard and Strauss, 1991; Devries *et al.*, 1989). Ku has also been reported to be capable of sequence-specific DNA interactions (Giffin *et al.*, 1996): for a review see (Dynan and Yoo, 1998). It has also been reported that Ku70 and possibly Ku80 are capable of interacting with DNA in the absence of their heterodimerisation partner (Chou *et al.*, 1992; Wang *et al.*, 1994). It has also been demonstrated that Ku heterodimers bound to DNA are able to specifically associate with one another (Cary *et al.*, 1997).

When complexed with DNA, Ku can interact with an approximately 460 kDa polypeptide, the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). DNA-PKcs is a member of the PI 3-kinase-like (PIKL) protein kinase family (Hartley *et al.*, 1995) and, together with Ku and DNA, forms a catalytically active DNA-PK complex with Ser/Thr kinase activity (Dvir *et al.*, 1992; Gottlieb and Jackson, 1993; Suwa *et al.*, 1994).

Little is currently known about how Ku interacts with DNA-

PKcs, although the fact that it is apparently unable to bind DNA-PKcs in the absence of DNA (Suwa et al., 1994) suggests that DNA binding by Ku induces a conformational change that permits the DNA-PKcs interaction. In particular, previous studies have provided little information about the regions of the Ku heterodimer which are involved in the DNA-PK_{cs} interaction.

It has been shown that, under certain conditions, DNA-PKcs
10 can bind to dsDNA ends and become activated in the absence of
Ku (Hammarsten and Chu, 1998; Yaneva et al., 1997). Thus,
whilst allosteric activation of DNA-PKcs by Ku might occur,
it appears that direct interactions between DNA-PKcs and DNA
can be sufficient to activate the kinase *in vitro*. This
15 suggests that contacts between DNA-PKcs and DNA play an
important role in DNA-PK activation, even in the presence of
Ku.

A breakthrough in the understanding of DNA-PKcs/Ku function
20 came with the discovery that defects in these proteins are
associated with a subset of mutant mammalian cell lines that
are defective in DNA DSB rejoining, and are profoundly
sensitive to ionising radiation and other agents that
generate DNA DSBs as their principal lethal lesion (Jackson,
25 S. P., et al (1995) TIBS 20, 412-415; Critchlow, S.E., et al
(1998) TIBS 23, 394-398). Indeed, the mutant phenotypes of
these cells are corrected by the introduction of the
appropriate Ku or DNA-PKcs expression vector, and recent work

using extracts of mammalian or *Xenopus laevis* cells has provided evidence for a direct involvement of Ku and DNA-PKcs in DNA DSB rejoining (Baumann and West, 1998; Labhart, P., (1999) Mol. Cell. Biol. 19, 2585-2593). Furthermore, DNA-PK catalytic activity has been implicated at an early stage of 5 DNA DSB repair in *Xenopus* cell-free extracts (Gu et al., 1996; Gu et al., 1998) and for radiation-induced DNA repair in cultured human cells (Okayasu, R., et al (1998) Radiat. Res. 149, 440-445). Coupled with the fact that Ku displays a 10 very high affinity for dsDNA ends *in vitro*, these data suggest that DNA-PK functions directly in the recognition and resolution of radiation-induced DNA DSBs *in vivo*.

Cells deficient in DNA-PKcs, Ku80, or Ku70 are also severely 15 impaired in V(D)J recombination, a site-specific genomic rearrangement process that takes place in the developing vertebrate immune system to help generate the vast antigen recognition capacity of antibody and T-cell receptor molecules (Jackson, S. P., et al (1995) TIBS 20, 412-415; Critchlow, S.E., et al (1998) TIBS 23, 394-398). This 20 process requires the production of DNA DSBs between the recombining gene segments by the RAG1/RAG2 proteins (Jackson, S. P., et al (1995) TIBS 20, 412-415; Critchlow, S.E., et al (1998) TIBS 23, 394-398) and the subsequent rejoining of the 25 DNA ends via DNA-PK-dependent mechanisms. For a single DNA rearrangement between two coding segments (V, D, or J regions) to occur, a join between the two coding sequences (known as the coding join) and one between the two non-coding

signal ends (the signal join) are made. Interestingly, Ku is essential for both types of join, whereas DNA-PKcs appears to be required only for coding joins and plays a non-essential and variable role in the generation of signal joins (Bogue et al., 1998). This suggests that, at least for the repair of a sub-set of DNA DSBs, Ku is able to function in the absence of DNA-PKcs.

Consistent with the Ku-associated DNA DSB repair pathway being highly conserved throughout eukaryotic evolution, Ku is found in *S. cerevisiae* and is essential for repair of DNA DSBs by the pathway of non homologous end-joining (Boulton and Jackson, 1996b; Siede et al., 1996 and see Critchlow and Jackson, 1998 for review). Perhaps surprisingly there is no clear orthologue of DNA-PKcs encoded by the fully-sequenced *S. cerevisiae* genome. Thus, in yeast, Ku carries out DNA-repair functions independently of DNA-PK. Although it is possible that the functions of mammalian DNA-PKcs are assumed by other members of the PIKL protein kinase family, such as Mecip and/or Telip, there is no evidence to suggest that these interact physically or genetically with Ku.

Interestingly, *S. cerevisiae* Ku has also been shown to play important roles in telomere length maintenance, and in the transcriptional silencing of genes placed close to telomeric DNA (Boulton and Jackson, 1996a; Boulton and Jackson, 1998 Porter et al., 1996).

The present inventors have investigated interactions between Ku70 and Ku80, and between the two Ku subunits and DNA-PKcs. The data presented herein lead to the conclusion that the two Ku subunits are structurally and functionally related to each 5 other, and appear to associate via a pseudo-homodimerisation mechanism. Furthermore, the work demonstrates that the extreme C-terminus of Ku80 plays an important role in the interaction between Ku and DNA-PKcs. These results provide for modulation of the structure and physiological functions 10 of DNA-PKcs and Ku, for instance by means of peptides corresponding to conserved regions in Ku70 or Ku80 and/or regions of interaction between Ku70 and Ku80 and/or other molecules such as DNA-PKcs, and allow for postulation of a model for the evolution of the DNA-PK complex.

15

Based on the experimental work and discussion herein the invention is further concerned with assays and methods for identifying homologues and orthologues of Ku polypeptides.

20 The present inventors have analysed the amino acid sequences of Ku70 and Ku80 polypeptides from a diverse range of species, and have identified six regions of sequence homology between them. These regions are referred to as Homology regions 1-6 (HRs 1-6) and are shown in Figure 1 as shaded 25 boxes. Corresponding regions from other sequences, e.g. from mammalian, e.g. rodent, e.g. mouse, hamster, sequences are easily identifiable by those skilled in the art using the work, discussion and Figures presented herein. References

to, and accession numbers of certain Ku70 and Ku80 sequences are provided in Table 1. Corresponding regions from the mouse sequences are provided in Table 2.

5 In one aspect the invention provides a peptide or polypeptide consisting of any of the amino acid sequences of any of the Homology Regions shown in Figure 1, or a peptide or polypeptide consisting of a variant of such a sequence. Variants are discussed elsewhere herein. Variants include

10 corresponding sequences from other animals, e.g. from mammals, e.g. from rodents. Examples of variants are the amino acid sequences of the Homology regions of mouse and hamster Ku70 or Ku80 sequences. The homology regions of the mouse sequences are provided in Table 2.

15 Variants of the sequences of the Homology Regions of human Ku70 and human Ku80 include the following: for Ku70 HR3, sequences which lack one or more residues corresponding to L420, D421, D422; for Ku80 HR3, sequences which lack one or more residues corresponding to N415, Y416, E417; for Ku70

20 HR5, sequences which lack one or more residues corresponding to T449, E450, K451, I452; for Ku80 HR5 sequences which lack one or more residues corresponding to S441, K442; for Ku70 HR6, sequences which lack one or more residues corresponding to E501, Q502; for Ku80 HR6 sequences which comprise one or

25 more residues corresponding to L505, P506.

A peptide or polypeptide consisting of any of the amino acid

sequences of any of the Homology Regions shown in Figure 1 or a peptide or polypeptide consisting of a variant of such a sequence, (a "Ku Homology Region Peptide"), may be used to establish antibodies which in turn may be used in the 5 identification of Ku polypeptides, or in the determination of the presence and/or quantity of Ku polypeptide in a test sample. Such antibodies form a further aspect of the invention in their own right. Methods of generating such antibodies, e.g. by raising them in a host organism or 10 selection from an immunoglobulin expression library, are discussed elsewhere herein, and are well known to those having skill in the art.

By way of example, antibodies which bind to peptides or 15 polypeptides consisting of any of the amino acid sequences of Ku Homology Region Peptide may be used to probe samples or expression libraries for Ku polypeptides, e.g. for the Ku polypeptide used in their generation, or for a cross-reactive Ku polypeptide. Binding of antibodies to Ku polypeptide may 20 be determined by any method known to those skilled in the art. Methods for determining binding and interaction are discussed herein. Samples may be tissue or cellular (e.g. nuclear) extracts, or *in vitro* translation systems expressing nucleic acids of interest. Expression libraries may be 25 bacterial or phage libraries.

Methods of determining the presence of, and optionally quantifying the amount of a Ku polypeptide in a test sample

may have a diagnostic purpose, e.g. in the diagnosis of any medical condition discussed herein (e.g. in a disease or disorder associated with loss of a Ku polypeptide). Such methods may also be used to evaluate a therapy to treat such 5 a condition.

Ku Homology Region Peptides as defined herein correspond to regions of Ku that are important for Ku function, e.g. for binding to proteins or nucleic acid, or for maintaining 10 tertiary structure. Accordingly, they may interact with other molecules and may be used to disrupt interaction between Ku and such other molecules. For instance a peptide corresponding to HR5 or HR6 may be used to interfere with interaction between Ku70 and Ku80, with functional or 15 biological consequences as discussed. Similarly, a peptide corresponding to any of HR1, HR2, HR3 or HR4 may be used to modulate a Ku-dependent activity or function which may have biological and/or therapeutic consequences.

20 In various further aspects, the present invention provides for nucleic acid encoding a polypeptide consisting of a Ku Homology Region Peptide, for vectors containing such nucleic acid, and for host cells containing such nucleic acid or such vectors. Nucleic acids, vectors and host cells are discussed 25 further elsewhere herein, and are useful in production of peptides and polypeptides by recombinant means.

Further peptides of interest in the present invention

10

correspond to the region of Ku80 identified herein as being necessary and sufficient for binding to DNA-PKcs, and may include or consist of the amino acid sequence EGGDVDDLLDMI.

Aspects of the invention based on this include assay methods for determining binding between Ku80 or EGGDVDDLLDMI on the one hand and DNA-PKcs on the other, identification of binding partners, and assay methods for agents that modulate, especially disrupt, such binding and which have therapeutic potential.

10

As used herein, the term "EGGDVDDLLDMI peptide" refers to a peptide which comprises or consists of the amino acid sequence EGGDVDDLLDMI and which has the ability to bind DNA-PK_{cs}. The term also refers to which comprise a variant of the sequence amino acid sequence EGGDVDDLLDMI and which have the ability to bind DNA-PK_{cs}. Variants are discussed elsewhere herein. Variants of the amino acid sequence EGGDVDDLLDMI include the corresponding regions of the related sequences shown in Figure 6 (multiple sequence alignment generated using the program "pileup" from the sequence analysis "Wisconsin Package, version 8.1" (Program Manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group: Gap creation 3.00; Gap extension 0.10).

EGGDVDDLLDMI peptides may for example be: (i) the amino acid sequence EGGDVDDLLDMI or a variant thereof; or (ii) the peptide of (i) fused to a heterologous amino acid sequence (see below), i.e. a sequence to which it is not naturally

fused in Ku80, or (iii) the peptide of (i) fused to a sequence with which it is naturally fused in wild-type Ku80, but in a peptide or polypeptide which has the ability to bind to DNA-PK_{cs}.

5

A peptide in accordance with any aspect of the present invention may include one or more heterologous amino acids joined to the specified peptide. By "heterologous" is meant not occurring in a Ku80 or Ku70 polypeptide joined by a peptide bond without intervening amino acids to the relevant specified peptide, that is to say usually a chain of amino acids which is not found naturally joined to the specified peptide at the position of fusion in the peptide of the invention. Usually where heterologous amino acids are included, the contiguous sequence of amino acids does not occur within Ku70 and/or Ku80, and may include or be 5 or more, preferably 10 or more, more preferably 15 or more, 20 or more or 30 or more amino acids with a sequence which does not occur contiguously in Ku70 and/or Ku80

20

A peptide or polypeptide according to the invention may be about 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-125, 125-150, 150-175, 175-200, 200-250, 250-300, 25 350-400, 400-450 amino acids in length. It may be about 22 amino acids or less, 69 amino acids or less, 137 amino acids or less, 208 amino acids or less, 259 amino acids or less, 314 amino acids or less, 394 amino acids or less in length.

A EGGDVDDLLDMI peptide of the invention is able to bind DNA-PK_{cs} and will be able to bind "DNA-PK_{cs} -like polypeptides", meaning not only the human DNA-dependent protein kinase catalytic subunit DNA-PK_{cs} and its homologues and orthologues in other species, for example in mouse, horse, or *Xenopus laevis*, but also variants of those DNA-PK_{cs} polypeptides, and polypeptides comprising an amino acid sequence which shares at least 30% amino acid sequence homology with the amino acid sequence of human DNA-PK_{cs} or a homologue or orthologue thereof, more preferably at least 35% sequence homology, more preferably at least 40% sequence homology, more preferably at least 50% sequence homology, more preferably at least 70% sequence homology, more preferably at least 80% sequence homology, still more preferably at least 90% sequence homology with the amino acid sequence of human DNA-PK_{cs} or a homologue or orthologue thereof. The amino acid sequence of human DNA-PK_{cs} is disclosed in Dynan and Yoo, 1998 and has the database accession number U47077. Sequence homology is defined elsewhere herein. The term DNA-PK_{cs}-like polypeptide may be used therefore to refer to related enzymes, e.g. other kinases of the PIKL family, e.g. ATM, ATR, FRAP. The term "DNA-PK_{cs}-like polypeptide" also encompasses alleles, mutants, derivatives and fragments of human DNA-PK_{cs} or a homologue or orthologue thereof.

25

Aspects of the present invention provide for the use of a peptide, whether a Ku Homology Region Peptide or an EGGDVDDLLDMI peptide or polypeptide, in screening or

searching for and/or obtaining/identifying a binding partner, such as for an EGGDVDDLLDMI peptide a DNA-PK_{cs}-like polypeptide which has the ability to bind to said EGGDVDDLLDMI.

5

Thus, a further aspect of the invention provides a screening or assay method for identifying an agent which binds to a peptide of the invention, or for determining binding of an peptide of the invention to an agent of interest may include:

10 (a) bringing a test substance into contact with said peptide; and
(b) determining binding of the test substance to said peptide.

15 In one embodiment, a screening or assay method for identifying an agent which binds to an EGGDVDDLLDMI peptide, or for determining binding of an EGGDVDDLLDMI peptide to an agent of interest may include:

20 (a) bringing a test substance into contact with said EGGDVDDLLDMI peptide; and
(b) determining binding of the test substance to said EGGDVDDLLDMI polypeptide.

25 A test substance which proves to be an agent which binds to an EGGDVDDLLDMI peptide may be, for example, a DNA-PKCS-like polypeptide or an antibody.

In other embodiments the peptide is a Ku Homology Region

Peptide, as disclosed.

The invention may be used to identify one or more regions of a polypeptide or other substance involved in an *in vivo* or an 5 *in vitro* interaction with a region of Ku70 or Ku80 corresponding to a Homology Region as disclosed or the sequence EGGDVDDLLDMI. This may involve identifying sequence motif(s) in a polypeptide which is/are involved in the interaction of polypeptide with the relevant amino acid 10 sequence in Ku80 or Ku70. Various fragments of a DNA-PK_{cs}-like polypeptide may be used in the above assays, e.g. fragments generated by N-terminal and/or C-terminal deletions of the full DNA-PK_{cs}-like polypeptide sequence.

15 Related aspects of the present invention provide the use of an peptide sequence of the invention for determining the presence in a test sample of a binding partner which has the ability to bind to said sequence, and the use of an agent which binds a peptide of the invention for determining the 20 presence in a test sample of a Ku70 or Ku80 peptide or polypeptide, such as an EGGDVDDLLDMI polypeptide.

A method for determining the presence in a test sample of an agent, e.g. polypeptide such as an antibody, which has the 25 ability to bind to a peptide of the invention, may include:

- (a) bringing a peptide of the invention into contact with the test sample; and
- (b) determining binding of the peptide to an agent if

present in the test sample.

A method for determining the presence in a test sample of a peptide of the invention, may include:

- 5 (a) bringing the an agent, e.g. polypeptide such as an antibody, into contact with the test sample; and
- (b) determining binding of the agent to a substance in the test sample.

10 A method for determining the presence in a test sample of any agent or substance may include quantifying the amount of the agent or substance in the sample.

15 Methods of determining the presence of an agent or substance in a test sample may have a diagnostic purpose, e.g. in the diagnosis of any medical condition discussed herein (e.g. in a disease or disorder associated with reduction or loss of a Ku-dependent function or biological activity). Such methods may also be used to evaluate a therapy to treat such a 20 condition.

Methods of determining binding and/or interaction in any method described herein are discussed below.

25 A test agent or substance employed in accordance with the present invention may be a natural or synthetic chemical compound.

A test agent or substance may be 0.005-0.01% pure, 0.01-0.05% pure, 0.05%-0.1% pure, 0.1-0.5% pure, 0.5-1% pure, 1-5% pure pure, 5-10% pure, 10-20% pure, 20-30% pure, 30-40% pure, 40-50% pure, 50-60% pure, 60-70% pure, 70-80% pure, 80-90% pure, 5 90-95% pure, or substantially pure.

A screening or assay method may include purifying and/or isolating a test substance and/or substance of interest from a mixture or extract, i.e. reducing the content of at least 10 one component of the mixture or extract, e.g. a component with which the test substance or substance of interest is naturally associated. The screening or assay method may include determining the ability of one or more fractions of a test mixture or extract to bind to peptide of the invention. 15 The purifying and/or isolating may employ any method known to those skilled in the art.

The precise format of any of the screening or assay methods of the present invention may be varied by those of skill in 20 the art using routine skill and knowledge. The skilled person is well aware of the need to employ appropriate control experiments.

A peptide of the invention and any agent identified by any 25 one of the methods provided by the present invention may be isolated and/or purified and/or further investigated and/or manufactured. Various methods and uses of such compounds are discussed elsewhere herein.

The inventors have identified a number of EGGDVDDLLDMI polypeptides which can bind to human DNA-PK_{cs}. These sequences form the basis of further aspects of the present invention.

5

Accordingly, the invention provides the EGGDVDDLLDMI polypeptides Ku80ΔN1, Ku80ΔN2, Ku80ΔN4, Ku80ΔN5, Ku80ΔN6, Ku80ΔN7, Ku80ΔN8 as shown in Figure 2 and variants thereof that retain ability to bind to a DNA-PK_{cs}-like polypeptide.

10

As used herein, variants of a stated amino acid sequence may have an amino acid sequence which shares at least about 30%, or 40%, or 50%, or 60%, or 70%, or 75%, or 80%, or 85%, 90% or 95% homology with the stated sequence. Homology is defined 15 elsewhere herein.

A variant may differ by one or more amino acid residues from the stated sequence, by one or more of addition, insertion, deletion and substitution of one or more amino acid residues. 20 It may include 1, 2, 3, 4, 5, or greater than 5 amino acid alterations such as substitutions with respect to the stated sequence.

A variant of a polypeptide for which the sequence is known or 25 disclosed herein may in certain embodiments be the same length or shorter than that sequence. In other embodiments the polypeptide (e.g. a DNA-PK_{cs}-like polypeptide or an EGGDVDDLLDMI polypeptide or a polypeptide consisting of any

of the amino acid sequences of any of the Homology Regions shown in Figure 1) may be included in a larger polypeptide. For example, 1, 2, 3, 4 or 5, 10, 20 or more additional amino acid residues, adjacent to a native form of the specific 5 polypeptide or heterologous thereto may be included at one end or both ends of the polypeptide.

Variants include mutants, alleles, derivatives and fragments of the stated sequence. Derivatives of polypeptides include 10 the polypeptide linked to a coupling partner, e.g. an effector molecule, a label, a drug, a toxin and/or a carrier or transport molecule, and/or a targeting molecule such as an antibody or binding fragment thereof or other ligand.

Techniques for coupling to both peptidyl and non-peptidyl 15 coupling partners are well known in the art. In one embodiment, the carrier molecule is a 16 amino acid peptide sequence derived from the homeodomain of *Antennapedia* (e.g. as sold under the name "Penetratin"), which can be coupled to a peptide via a terminal Cys residue. The "Penetratin" 20 molecule and its properties are described in WO 91/18981.

As is well-understood, homology at the amino acid level is generally in terms of amino acid similarity or identity. Similarity allows for "conservative variation", i.e. 25 substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine

for asparagine. Homology may be taken over the full-length of a sequence or over a part, such as 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 200 contiguous nucleotides or amino acids. That two nucleotide sequences are said to share 5 "homology" or be "homologous" is based on sequence comparison. Any phylogenetic relationship is irrelevant for this. Those skilled in the art routinely refer to homology between nucleotide sequences with no implication for evolutionary origin. Two homologous nucleotide sequences may 10 also be said to be "similar" or have a certain per centage similarity or a certain per centage identity.

In general it is not critical which of the various standard algorithms are used to determine how homologous two 15 nucleotide sequences are with one another. A preferred algorithm may be GAP, which uses the alignment method of Needleman and Wunsch (*J. Mol. Biol.* (1970) 48, 443-453) and is included in the Program Manual or the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, 575 20 Science Drive, Madison, Wisconsin, USA). In the absence of instructions to the contrary, the skilled person would understand to use the default parameters with the aim of maximizing alignment, with a gap creation penalty = 12 and gap extension penalty = 4.

25

Similarity or homology (the terms are used interchangeably) or identity may be as defined and determined by the TBLASTN program, of Altschul et al. (1990) *J. Mol. Biol.* 215: 403-10,

or BestFit, which is part of the Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA, Wisconsin 53711). Preferably sequence comparisons are made using FASTA and FASTP (see 5 Pearson & Lipman, 1988. Methods in Enzymology 183: 63-98). Parameters are preferably set, using the default matrix, as follows: Gapopen (penalty for the first residue in a gap): - 12 for proteins / -16 for DNA; Gapext (penalty for additional residues in a gap): -2 for proteins / -4 for DNA; KTUP word 10 length: 2 for proteins / 6 for DNA.

Nucleic acid sequence homology may be determined by means of selective hybridisation between molecules under stringent conditions.

15 Preliminary experiments may be performed by hybridising under low stringency conditions. For probing, preferred conditions are those which are stringent enough for there to be a simple pattern with a small number of hybridisations identified as 20 positive which can be investigated further.

For example, hybridizations may be performed, according to the method of Sambrook et al. (below) using a hybridization solution comprising: 5X SSC (wherein 'SSC' = 0.15 M sodium 25 chloride; 0.15 M sodium citrate; pH 7), 5X Denhardt's reagent, 0.5-1.0% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at

least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes - 1 hour at 37°C in 1X SSC and 1% SDS; 5 (4) 2 hours at 42-65°C in 1X SSC and 1% SDS, changing the solution every 30 minutes.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid 10 molecules of a specified sequence homology is (Sambrook et al., 1989): $T_m = 81.5^\circ\text{C} + 16.6\log [\text{Na}^+] + 0.41 (\% \text{ G+C}) - 0.63 (\% \text{ formamide}) - 600/\#\text{bp}$ in duplex.

As an illustration of the above formula, using $[\text{Na}^+] = 15 [0.368]$ and 50-% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence 20 identity would be observed using a hybridization temperature of 42°C. Such a sequence would be considered substantially homologous to the nucleic acid sequence of the present invention.

It is well known in the art to increase stringency of 25 hybridisation gradually until only a few positive clones remain. Other suitable conditions include, e.g. for detection of sequences that are about 80-90% identical, hybridization overnight at 42°C in 0.25M Na_2HPO_4 , pH 7.2, 6.5% SDS, 10%

dextran sulfate and a final wash at 55°C in 0.1X SSC, 0.1% SDS. For detection of sequences that are greater than about 90% identical, suitable conditions include hybridization overnight at 65°C in 0.25M Na₂HPO₄, pH 7.2, 6.5% SDS, 10% 5 dextran sulfate and a final wash at 60°C in 0.1X SSC, 0.1% SDS.

Polypeptides in accordance with and for use in the present invention, e.g. EGGDVDDLLDMI polypeptides or DNA-PK_{cs}-like 10 polypeptides or peptides or polypeptides consisting of a Ku Homology Region peptide as defined, may be generated wholly or partly by chemical synthesis, in accordance with well-established techniques, such as standard liquid or, preferably, solid-phase peptide synthesis methods, general 15 descriptions of which are broadly available (see, for example, in J.M. Stewart and J.D. Young, Solid Phase Peptide Synthesis, 2nd edition, Pierce Chemical Company, Rockford, Illinois (1984), in M. Bodanzsky and A. Bodanzsky, The Practice of Peptide Synthesis, Springer Verlag, New York 20 (1984); and Applied Biosystems 430A Users Manual, ABI Inc., Foster City, California), or they may be prepared in solution, by the liquid phase method or by any combination of solid-phase, liquid phase and solution chemistry.

25 Another convenient way of producing peptides and polypeptides according to or for use in the present invention is to express encoding nucleic acid, by use of nucleic acid in an expression system.

Nucleic acid molecules encoding a peptide or polypeptide as defined herein in accordance with the present invention represent further aspects of the present invention in their own right.

5

Generally, nucleic acid according to, or for use in, the present invention is provided as an isolate, in isolated and/or purified form, or free or substantially free of material with which it is naturally associated, such as free or substantially free of nucleic acid flanking the gene in the (e.g. human) genome, except possibly one or more regulatory sequence(s) for expression. Nucleic acid may be wholly or partially synthetic and may include genomic DNA, cDNA or RNA.

15

Nucleic acid sequences encoding a peptide or polypeptide according to the invention, may be readily prepared by the skilled person using the information and references contained herein and techniques known in the art (for example, see 20 Sambrook, Fritsch and Maniatis, "Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, and Ausubel et al, Short Protocols in Molecular Biology, John Wiley and Sons, 1992). These techniques include (i) the use 25 of the polymerase chain reaction (PCR) to amplify samples of such nucleic acid, e.g. from genomic sources, (ii) chemical synthesis, or (iii) preparing cDNA sequences.

DNA encoding a peptide or polypeptide according to the

invention may be generated and used in any suitable way known to those of skill in the art, including taking encoding DNA, identifying suitable restriction enzyme recognition sites either side of the portion to be expressed, and cutting out 5 said portion from the DNA. The portion may then be operably linked to a suitable promoter in a standard commercially available expression system. Another recombinant approach is to amplify the relevant portion of the DNA with suitable PCR primers.

10

Modifications to a nucleic acid sequence may be made, e.g. using site directed mutagenesis, to lead to the production of modified forms of a polypeptide, e.g. an EGGDVDDLLDMI polypeptide, e.g. a mutant form of such a sequence, or to 15 take account of codon preference in the host cells used to express the nucleic acid.

In order to obtain expression of the nucleic acid sequences of the invention, the sequences may be incorporated in a 20 vector having one or more control sequences operably linked to the nucleic acid to control its expression. Vectors may contain appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as 25 appropriate, e.g. nucleic acid sequences so that the polypeptide or peptide is produced as a fusion and/or nucleic acid encoding secretion signals so that the polypeptide produced in the host cell is secreted from the cell. Vectors

may be plasmids, viral e.g. phage, or phagemid, as appropriate. For further details see, for example, Molecular Cloning: a Laboratory Manual: 2nd edition, Sambrook et al., 1989, Cold Spring Harbor Laboratory Press. Many known 5 techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in Current Protocols in Molecular Biology, Ausubel et 10 al. eds., John Wiley & Sons, 1992.

A peptide or polypeptide according to the invention may then be obtained by transforming the vectors into host cells in which the vector is functional, culturing the host cells so 15 that the encoded peptide or polypeptide, e.g. EGGDVDDLLDMI polypeptide, is produced and recovering the sequence from the host cells or the surrounding medium.

A further aspect of the present invention provides a host 20 cell containing heterologous nucleic acid as disclosed herein.

Systems for cloning and expression of polypeptides in a variety of different host cells are well known. Suitable 25 host cells include bacteria, eukaryotic cells such as mammalian and yeast, and baculovirus systems. Mammalian cell lines available in the art for expression of a heterologous polypeptide include Chinese hamster ovary cells, HeLa cells,

baby hamster kidney cells, COS cells and many others. A common, preferred bacterial host is *E. coli*.

The nucleic acid of the invention may be integrated into the 5 genome (e.g. chromosome) of the host cell. Integration may be promoted by inclusion of sequences which promote recombination with the genome, in accordance with standard techniques. The nucleic acid may be on an extra-chromosomal vector within the cell, or otherwise identifiably 10 heterologous or foreign to the cell.

Thus, a host cell containing nucleic acid according to the present invention, e.g. as a result of introduction of the nucleic acid into the cell or into an ancestor of the 15 cell (which introduction may take place *in vivo* or *ex vivo*), may be comprised (e.g. in the soma) within an organism which is an animal, particularly a mammal, which may be human or non-human, such as rabbit, guinea pig, rat, mouse or other rodent, cat, dog, pig, sheep, goat, cattle or horse, or a 20 bird, such as a chicken.

Genetically modified or transgenic animals or birds comprising such a cell are also provided as further aspects of the present invention.

This may have a therapeutic aim. (Gene therapy is discussed elsewhere herein). Also, the presence of a mutant, allele, derivative or other variant sequence within cells of an

organism may allow the organism to be used as a model in studying substances which modulate binding of a peptide to a binding partner, e.g. an EGGDVDDLLDMI polypeptide to a DNA-PK_{cs}-like polypeptide. Conveniently, however, at least 5 preliminary assays for such substances may be carried out *in vitro*, that is within host cells or in cell-free systems. Where an effect of a test compound is established on cells *in vitro*, those cells or cells of the same or similar type may be grafted into an appropriate host animal for *in vivo* 10 testing.

A further aspect provides a method which includes introducing a nucleic acid molecule of the invention into a host cell. The introduction, which may (particularly for *in vitro* 15 introduction) be generally referred to without limitation as "transformation", may employ any available technique. For eukaryotic cells, suitable techniques may include calcium phosphate transfection, DEAE-Dextran, electroporation, liposome-mediated transfection and transduction using 20 retrovirus or other virus, e.g. vaccinia or, for insect cells, baculovirus. For bacterial cells, suitable techniques may include calcium chloride transformation, electroporation and transfection using bacteriophage. As an alternative, direct injection of the nucleic acid could be employed.

25

Marker genes such as antibiotic resistance or sensitivity genes may be used in identifying clones containing nucleic acid of interest, as is well known in the art.

The introduction may be followed by causing or allowing expression from the nucleic acid, e.g. by culturing host cells (which may include cells actually transformed although more likely the cells will be descendants of the transformed cells) under conditions for expression of the gene, so that the encoded polypeptide is produced. If the polypeptide is expressed coupled to an appropriate signal leader peptide it may be secreted from the cell into the culture medium.

Following production by expression, a polypeptide may be isolated and/or purified from the host cell and/or culture medium, as the case may be, and subsequently used as desired, e.g. in the formulation of a composition which may include one or more additional components, such as a pharmaceutical composition which includes one or more pharmaceutically acceptable excipients, vehicles or carriers (e.g. see below).

Introduction into target cells of nucleic acid of the present invention may take place *in vivo* by way of gene therapy, for instance to modulate e.g. disrupt or interfere with, binding of an EGGDVDDLLDMI polypeptide to a DNA-PK_{cs}-like polypeptide.

In light of the above, the present invention also provides a method of making a peptide or polypeptide as defined, the method including expression from nucleic acid encoding the peptide or polypeptide. This may conveniently be achieved by growing a host cell containing the nucleic acid in culture under appropriate conditions which cause or allow expression of the peptide or polypeptide. Note however that expression

may also be carried out in *in vitro* systems, e.g. reticulocyte lysate, as discussed elsewhere herein.

As discussed above, the present invention provides
5 EGGDVDDLLDMI polypeptides, methods of identifying further
EGGDVDDLLDMI polypeptides which have the ability to bind to a
DNA-PK_{cs}-like polypeptide and methods of identifying DNA-PK_{cs}-
like polypeptides which have the ability to bind to an
EGGDVDDLLDMI polypeptide.

10

EGGDVDDLLDMI polypeptides having the ability to bind to a
DNA-PK_{cs}-like polypeptide will have the ability to modulate,
e.g. interfere with, binding of that DNA-PK_{cs}-like polypeptide
to proteins harbouring that EGGDVDDLLDMI polypeptide or
15 variants thereof, e.g. to proteins which are *in vivo* binding
partners of that DNA-PK_{cs}-like polypeptide.

Accordingly, in still further aspects, the present invention
provides for assays and methods of screening for a substance
20 which modulates, e.g. inhibits, reduces or interferes with,
binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI
polypeptide, and for the use of DNA-PK_{cs}-like polypeptides
and/or EGGDVDDLLDMI polypeptides in these screening methods
and assays.

25

Such a method may include:

(i) bringing a DNA-PK_{cs}-like polypeptide into contact with an
EGGDVDDLLDMI polypeptide in the presence of one or more test

substances;

(ii) determining binding of the DNA-PK_{cs}-like polypeptide to the EGGDVDDLLDMI polypeptide.

5 Binding in the presence of a test substance may be compared with binding of the DNA-PK_{cs}-like polypeptide to the EGGDVDDLLDMI polypeptide in comparable reaction medium and conditions in the absence of a test substance. A test substance able to modulate the binding may be identified.

10

Agents which modulate e.g. increase or potentiate interaction between a DNA-PK_{cs}-like polypeptide and an EGGDVDDLLDMI polypeptide may be identified using conditions which, in the absence of a positively-testing agent, prevent or impair the 15 substances interacting.

A difference in the binding of the DNA-PK_{cs}-like polypeptide to the EGGDVDDLLDMI polypeptide between the treated and untreated conditions is indicative of a modulating effect of 20 the relevant test substance(s).

Further aspects of the invention provide assay methods for identifying or obtaining agents which modulate, especially interfere with, interaction between Ku70 or Ku80 and a 25 binding partner, via a Homology Region as identified herein. Such methods may comprise

(i) bringing a Ku Homology Region Peptide into contact with a binding partner in the presence of one or more test

substances;

(ii) determining binding of the Ku Homology Region Peptide to the binding partner.

5 Binding may be determined by any suitable method known to those skilled in the art. Examples of such methods are described below.

In any assay method according to the invention, the amount of 10 test substance or compound which may be added to an assay of the invention will normally be determined by trial and error depending upon the type of compound used. Typically, from about 0.001 nM to 1mM or more concentrations of putative inhibitor compound may be used, for example from 0.01 nM to 15 100 μ M, e.g. 0.1 to 50 μ M, such as about 10 μ M. Greater concentrations may be used when a peptide is the test substance. Even a molecule which has a weak effect may be a useful lead compound for further investigation and development.

20

Compounds which may be screened may be natural or synthetic chemical compounds used in drug screening programmes.

Extracts of plants, microbes or other organisms, which contain several characterised or uncharacterised components 25 may also be used.

It is worth noting that combinatorial library technology provides an efficient way of testing a potentially vast

number of different substances for ability to modulate an interaction. Such libraries and their use are known in the art, for all manner of natural products, small molecules and peptides, among others. The use of peptide libraries may be
5 preferred in certain circumstances.

In various aspects the present invention provides a modulator identified by a screening method of the invention, e.g. a substance which interferes with or interrupts, increases or
10 potentiates binding of a DNA-PK_{cs}-like polypeptide (e.g. DNA-PK_{cs}) to a target EGGDVDDLLDMI polypeptide (e.g. Ku80 or an EGGDVDDLLDMI polypeptide of the present invention), or Ku70 and/or Ku80 to each other or a binding partner (via a Homology Region).
15

Following identification of a modulator, the substance may be purified and/or investigated further and/or manufactured. A modulator may be used to obtain peptidyl or non-peptidyl mimetics, e.g. by methods well known to those skilled in the
20 art and discussed herein. It may be used in a therapeutic context as discussed below.

One class of modulators comprises peptides comprising the amino acid sequence EGGDVDDLLDMI or variants thereof, which
25 peptides have the ability to bind to a DNA-PK_{cs}-like polypeptide. A further class of modulators comprises peptide fragments of DNA-PK_{cs}-like polypeptides, or variants thereof, particularly fragments of DNA-PK_{cs}-like polypeptides which

conatin sequence motifs that have been identified (e.g. in a screening method of the present invention) as being involved in binding of the DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI sequence.

5

Suitable peptide modulators are those which bind to Ku80 or to an EGGDVDDLLDMI polypeptide shown in Figure 4 and/or which have a length of 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-85, 85-90, 90-95, 95-100, or more than 100 amino acids.

10 Nucleic acid encoding such peptide frgaments, vectors and host cells containing such nucleic acid, and methods of expressing nucleic acid encoding such fragments are further aspects of the present invention. The discussion included herein relating to the construction, use and expression of
15 peptides and polypeptides and encoding nucleic acid applies to such embodiments as appropriate.

Antibodies directed to the a peptide disclosed herein, a Ku Homology Region peptide or EGGDVDDLLDMI, or to a sequence in
20 a DNA-PK_{cs}-like polypeptide identified (e.g. in a method of the present invention) as being involved in an interaction with an EGGDVDDLLDMI polypeptide, form further classes of putative inhibitor compounds and represent individual aspects of the invention in their own right. Candidate inhibitor
25 antibodies may be characterised and their binding regions determined to provide single chain antibodies and fragments thereof which are responsible for disrupting the binding.

Antibodies in accordance with and/or for use in the present invention may be obtained using techniques which are standard in the art. Methods of producing antibodies include immunising a mammal (e.g. mouse, rat, rabbit, horse, goat, sheep or monkey) with a polypeptide or fragment thereof, e.g. with a EGGDVDDLLDMI polypeptide (e.g. with the polypeptide EGGDVDDLLDMI or a variant thereof), or with a DNA-PK_{cs}-like polypeptide, or with a Ku Homology Region Peptide.

Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and screened, preferably using binding of antibody to antigen of interest. For instance, Western blotting techniques or immunoprecipitation may be used (Armitage et al., 1992, Nature 357: 80-82). Isolation of antibodies and/or antibody-producing cells from an animal may be accompanied by a step 15 of sacrificing the animal.

As an alternative or supplement to immunising a mammal with a peptide, an antibody specific for a peptide or polypeptide of interest may be obtained from a recombinantly produced library of expressed immunoglobulin variable domains, e.g. using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047. The library may be 20 naïve, that is constructed from sequences obtained from an organism which has not been immunised with the relevant peptide or polypeptide (or fragments thereof), or may be one constructed using sequences obtained from an organism which 25

has been exposed to the antigen of interest.

Antibodies according to and/or for use in the present invention may be modified in a number of ways. Indeed the 5 term "antibody" should be construed as covering any binding substance having a binding domain with the required specificity. Thus the invention covers antibody fragments, derivatives, functional equivalents and homologues of antibodies, including synthetic molecules and molecules whose 10 shape mimicks that of an antibody enabling it to bind an antigen or epitope.

Example antibody fragments, capable of binding an antigen or other binding partner are the Fab fragment consisting of the 15 VL, VH, Cl and CH1 domains; the Fd fragment consisting of the VH and CH1 domains; the Fv fragment consisting of the VL and VH domains of a single arm of an antibody; the dAb fragment which consists of a VH domain; isolated CDR regions and F(ab')2 fragments, a bivalent fragment including two Fab 20 fragments linked by a disulphide bridge at the hinge region. Single chain Fv fragments are also included.

A hybridoma producing a monoclonal antibody according to the present invention may be subject to genetic mutation or other 25 changes. It will further be understood by those skilled in the art that a monoclonal antibody can be subjected to the techniques of recombinant DNA technology to produce other antibodies or chimeric molecules which retain the specificity

of the original antibody. Such techniques may involve introducing DNA encoding the immunoglobulin variable region, or the complementarity determining regions (CDRs), of an antibody to the constant regions, or constant regions plus 5 framework regions, of a different immunoglobulin. See, for instance, EP184187A, GB 2188638A or EP-A-0239400. Cloning and expression of chimeric antibodies are described in EP-A-0120694 and EP-A-0125023.

10 Hybridomas capable of producing antibody with desired binding characteristics are within the scope of the present invention, as are host cells, eukaryotic or prokaryotic, containing nucleic acid encoding antibodies (including antibody fragments) and capable of their expression. The 15 invention also provides methods of production of the antibodies including growing a cell capable of producing the antibody under conditions in which the antibody is produced, and preferably secreted.

20 The reactivities of antibodies on a sample may be determined by any appropriate means. Tagging with individual reporter molecules is one possibility. The reporter molecules may directly or indirectly generate detectable, and preferably measurable, signals. The linkage of reporter molecules may 25 be directly or indirectly, covalently, e.g. via a peptide bond or non-covalently. Linkage via a peptide bond may be as a result of recombinant expression of a gene fusion encoding antibody and reporter molecule. The mode of determining

binding is not a feature of the present invention and those skilled in the art are able to choose a suitable mode according to their preference and general knowledge.

- 5 Antibodies may also be used in purifying and/or isolating a polypeptide, e.g. an EGGDVDDLLDMI polypeptide or a Ku polypeptide or DNA-PK_{cs}-like polypeptide for instance following production of the polypeptide by expression from encoding nucleic acid therefor. The antibodies may be
- 10 respectively directed to the amino acid sequence EGGDVDDLLDMI, to a Ku Homology Region Peptide, or to a region of a DNA-PK_{cs}-like polypeptide identified (e.g. by a method of the present invention) as having the ability to interact with an EGGDVDDLLDMI polypeptide. Antibodies may be useful in a
- 15 therapeutic context (which may include prophylaxis) e.g. to disrupt binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide with a view to inhibiting the activity of either component. Antibodies can for instance be micro-injected into cells, e.g. at a tumour site, subject to
- 20 radio- and/or chemo-therapy (as discussed already above). Antibodies may be employed in accordance with the present invention for other therapeutic and non-therapeutic purposes which are discussed elsewhere herein.
- 25 In a further aspect, the present invention provides the use of a peptide as disclosed herein in a method of designing a peptide or non-peptidyl mimetic of the polypeptide. A mimetic may be able to bind to a DNA-PK_{cs}-like polypeptide

and/or modulate interaction between a DNA-PK_{cs}-like polypeptide and an EGGDVDDLLDMI polypeptide. A EGGDVDDLLDMI polypeptide used in such a method may be the amino acid sequence EGGDVDDLLDMI or a variant thereof, or an 5 EGGDVDDLLDMI polypeptide of the present invention, e.g. one shown in Figure 4.

The present invention similarly provides for the use of a DNA-PK_{cs}-like polypeptide, for example DNA-PK_{cs}, particularly a 10 DNA-PK_{cs}-like polypeptide identified as having the ability to bind to the amino acid sequence EGGDVDDLLDMI, in a method of designing a peptide or non-peptidyl mimetic of a DNA-PK_{cs}-like polypeptide, which mimetic is able to bind to a given EGGDVDDLLDMI polypeptide, e.g. to Ku80.

15

Accordingly, the present invention provides a method of designing a mimetic of a EGGDVDDLLDMI polypeptide which has the biological activity of binding to a DNA-PK_{cs}-like polypeptide (e.g. DNA-PK_{cs}), or a method of designing a 20 mimetic of a DNA-PK_{cs}-like polypeptide which has biological activity of binding to a target EGGDVDDLLDMI polypeptide (e.g. to Ku80 and/or to an EGGDVDDLLDMI polypeptide of the present invention), or a method of designing a mimetic of a Ku Homology Region Peptide, said method comprising:

25 (i) analysing a substance having the biological activity to determine the amino acid residues essential and important for the activity to define a pharmacophore; and, (ii) modelling the pharmacophore to design and/or screen

candidate mimetics having the biological activity.

Suitable modelling techniques are known in the art. This includes the study of the bonding between peptides or 5 polypeptides and to design compounds which contain functional groups arranged in such a manner that they could reproduce that bonding.

The designing of mimetics to a known pharmaceutically active 10 compound is a known approach to the development of pharmaceuticals based on a "lead" compound. This might be desirable where the active compound is difficult or expensive to synthesise or where it is unsuitable for a particular method of administration, for instance polypeptides of the 15 invention may not be well suited as active agents for oral compositions as they tend to be quickly degraded by proteases in the alimentary canal.

There are several steps commonly taken in the design of a 20 mimetic from a compound having a given target property. Firstly, the particular parts of the compound that are critical and/or important in determining the target property are determined. In the case of a peptide, this can be done by systematically varying the amino acid residues in the 25 peptide, e.g. by substituting each residue in turn. These parts or residues constituting the active region of the compound are known as its "pharmacophore".

Once the pharmacophore has been found, its structure is modelled according to its physical properties, e.g. stereochemistry, bonding, size and/or charge, using data from a range of sources, e.g. spectroscopic techniques, X-ray 5 diffraction data and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of a pharmacophore, rather than the bonding between atoms) and other techniques can be used in this modelling process.

10 In a variant of the above approach, the three-dimensional structure of a ligand and its binding partner are modelled. This can be especially useful where the ligand and/or binding partner change conformation on binding, allowing the model to take account of this in the design of the mimetic.

15 A template molecule is then selected onto which chemical groups which mimic the pharmacophore can be grafted. The template molecule and the chemical groups grafted on to it can conveniently be selected so that the mimetic is easy to 20 synthesise, is likely to be pharmacologically acceptable, and does not degrade *in vivo*, while retaining the biological activity of the lead compound. The mimetic or mimetics found by this approach can then be screened to see whether they have the target property, or to what extent they exhibit it.

25 Further optimisation or modification can then be carried out to arrive at one or more final mimetics for *in vivo* or clinical testing.

The mimetic or mimetics found by any of the approaches described herein may be used in the assay methods of the present invention to determine whether they have the ability to bind to the relevant compound, e.g. to a DNA-PK_{cs}-like polypeptide (e.g. DNA-PK_{cs}) or to a target EGGDVDDLLDMI polypeptide (e.g. to Ku80 and/or to an EGGDVDDLLDMI polypeptide of the present invention).

Mimetics obtained by a method of the invention form a further aspect of the invention.

The invention further provides various therapeutic methods and uses of one or more substances selected from (i) an EGGDVDDLLDMI polypeptide, (e.g. an EGGDVDDLLDMI polypeptide according to the present invention, a fragment of Ku80 comprising the sequence EGGDVDDLLDMI, an EGGDVDDLLDMI polypeptide identified in accordance with a method of the present invention); (ii) a DNA-PK_{cs}-like polypeptide, (e.g. a variant of DNA-PK_{cs}, e.g. fragment of DNA-PK_{cs}, a DNA-PK_{cs}-like polypeptide identified by a screening method of the present invention); (iii) any peptide or polypeptide disclosed herein, e.g. a Ku Homology Region Peptide; (iv) a modulator identified by a screening method of the present invention; (v) a mimetic of any of the above substances which can bind to a DNA-PK_{cs}-like polypeptide or EGGDVDDLLDMI polypeptide.

The therapeutic/prophylactic purpose of such a method or use may be the modulation, e.g. disruption or interference, of

the binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide, or a Ku Homology Region to a binding partner, e.g. to modulate any activity mediated by virtue of such binding.

5

The therapeutic/prophylactic purpose may for example be:

- (i) Chemotherapy and/or radiotherapy, e.g. sensitising tumours or other cells to radiation or chemotherapy, e.g. modulating (for example inhibiting) interactions of Ku70, Ku80 and/or DNA-PK_{cs} leading to e.g. impairment of tumour proliferation and/or growth;
- (ii) Modulation (e.g. inhibition) of infection of pathogens (e.g. viruses) whose infectivity is influenced by Ku and/or DNA-PK_{cs}, e.g. inhibition of retroviral integration;
- (iii) Controlling cell e.g. tumour growth via modulation of telomere function;
- (iv) Modulating V(D)J recombination or other genomic rearrangement that employs Ku and/or DNA-PK_{cs}.

20 In various further aspects the present invention thus provides a pharmaceutical composition, medicament, drug or other composition for such a purpose, the composition comprising one or more of the substances set out above, the

25

use of such a substance in a method of medical treatment, a method comprising administration of such a substance or composition to a patient, e.g. for treatment (which may include preventative treatment) of a medical condition, e.g.

5 a condition associated with a defect or disorder in DNA repair, or cell cycle control, e.g. for treatment of a disorder of cellular proliferation such as cancer, use of such a substance in the manufacture of a composition, medicament or drug for administration for such a purpose, 10 e.g. for treatment of a proliferative disorder, and a method of making a pharmaceutical composition comprising admixing such a substance with a pharmaceutically acceptable excipient, vehicle or carrier, and optionally other ingredients.

15

The substances may be used as sole active agents or in combination with one another or with any other active substance, e.g. for anti-tumour therapy another anti-tumour compound or therapy, such as radiotherapy or chemotherapy.

20

Whatever the substance used in a method of medical treatment of the present invention, administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis 25 may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated.

Prescription of treatment, e.g. decisions on dosage etc, is within the responsibility of general practitioners and other medical doctors.

- 5 A substance or composition may be administered alone or in combination with other treatments, either simultaneously or sequentially dependent upon the condition to be treated, e.g. cancer.
- 10 Pharmaceutical compositions according to the present invention, and for use in accordance with the present invention, may include, in addition to active ingredient, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material will depend on the route of administration, which may be oral, or by injection, e.g. cutaneous, subcutaneous or intravenous.
- 20 Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may include a solid carrier such as gelatin or an adjuvant. Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may

be included.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient 5 will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, 10 Ringer's Injection, Lactated Ringer's Injection.

Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

Liposomes, particularly cationic liposomes, may be used in 15 carrier formulations.

Examples of techniques and protocols mentioned above can be found in Remington's Pharmaceutical Sciences, 16th edition, Osol, A. (ed), 1980.

20

The substance or composition may be administered in a localised manner to a tumour site or other desired site or may be delivered in a manner in which it targets tumour or other cells.

25

Targeting therapies may be used to deliver the active substance more specifically to certain types of cell, by the use of targeting systems such as antibody or cell specific

ligands. Targeting may be desirable for a variety of reasons, for example if the agent is unacceptably toxic, or if it would otherwise require too high a dosage, or if it would not otherwise be able to enter the target cells.

5

Instead of administering such substances directly, they may be produced in the target cells by expression from an encoding nucleic acid introduced into the cells, e.g. from a viral vector. The vector may be targeted to the specific 10 cells to be treated, or it may contain regulatory elements which are switched on more or less selectively by the target cells.

Nucleic acid encoding the substance e.g. a polypeptide able 15 to modulate, e.g. interfere with, the binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide may thus be used in methods of gene therapy, for instance in treatment of individuals, e.g. with the aim of preventing or curing (wholly or partially) a disorder.

20

Vectors such as viral vectors have been used in the prior art to introduce nucleic acid into a wide variety of different target cells. Typically the vectors are exposed to the target cells so that transfection can take place in a sufficient proportion of the cells to provide a useful 25 therapeutic or prophylactic effect from the expression of the desired polypeptide. The transfected nucleic acid may be permanently incorporated into the genome of each of the

targeted cells, providing long lasting effect, or alternatively the treatment may have to be repeated periodically.

5 A variety of vectors, both viral vectors and plasmid vectors, are known in the art, see US Patent No. 5,252,479 and WO 93/07282. In particular, a number of viruses have been used as gene transfer vectors, including papovaviruses, such as SV40, vaccinia virus, herpesviruses, including HSV and EBV, 10 and retroviruses. Many gene therapy protocols in the prior art have used disabled murine retroviruses.

As an alternative to the use of viral vectors in gene therapy other known methods of introducing nucleic acid into cells 15 includes mechanical techniques such as microinjection, transfer mediated by liposomes and receptor-mediated DNA transfer.

Receptor-mediated gene transfer, in which the nucleic acid is 20 linked to a protein ligand via polylysine, with the ligand being specific for a receptor present on the surface of the target cells, is an example of a technique for specifically targeting nucleic acid to particular cells.

25 A peptide, polypeptide or other substance according to the present invention, e.g. a nucleic acid molecule which encodes a peptide or polypeptide, may be provided in a kit, e.g. sealed in a suitable container which protects its contents

from the external environment. Such a kit may include instructions for use.

In still further aspects the present invention provides for 5 the purification of a DNA-PK_{cs}-like polypeptide or protein containing a DNA-PK_{cs}-like polypeptide, or for the purification of an EGGDVDDLLDMI polypeptide. The invention also provides for a purified DNA-PK_{cs}-like polypeptide or protein containing a DNA-PK_{cs}-like polypeptide and a purified 10 EGGDVDDLLDMI polypeptide. The purified protein or polypeptide may be about 10% pure, more preferably about 20% pure, more preferably about 30% pure, more preferably about 40% pure, more preferably about 50% pure, more preferably about 60% pure, more preferably about 70% pure, more 15 preferably about 80% pure, more preferably about 90% pure, more preferably about 95% pure, or substantially pure.

The present invention thus provides a method of purifying a DNA-PK_{cs}-like polypeptide or protein containing a DNA-PK_{cs}-like 20 polypeptide, the method including contacting the DNA-PK_{cs}-like polypeptide with an EGGDVDDLLDMI peptide. The present inventors have already shown that the DNA-PK_{cs}-like polypeptide human DNA-PK_{cs} can be purified to virtual 25 homogeneity in a single step using a peptide of amino acid sequence EGGDVDDLLDMI.

A mixture of material including a DNA-PK_{cs}-like polypeptide or protein containing a DNA-PK_{cs}-like polypeptide may be

contacted against immobilised EGGDVDDLLDMI polypeptide (e.g. immobilised either covalently or non-covalently such as via a specific binding molecule such as streptavidin or biotin) and molecules which do not bind to the phosphopeptide are washed off.

Likewise, the invention provides a method of purifying an EGGDVDDLLDMI polypeptide, the method including contacting material containing the polypeptide with a DNA-PK_{cs}-like polypeptide.

Preferred EGGDVDDLLDMI polypeptides and DNA-PK_{cs}-like polypeptides for use in methods of the invention are discussed elsewhere herein.

The DNA-PK_{cs}-like polypeptide or protein containing a DNA-PK_{cs}-like polypeptide which is contacted with the EGGDVDDLLDMI polypeptide, or the EGGDVDDLLDMI polypeptide which is contacted with the DNA-PK_{cs}-like polypeptide, in a purification method of the present invention, may be in a mixture of molecules, such as a cellular extract, such as a normal cell of an organism such as a human or a recombinant host cell expressing the protein or polypeptide from encoding DNA, such as a bacterial, eukaryotic (e.g. mammalian or yeast) or insect cell, such as in a baculovirus expression system. Purification may follow production of such a polypeptide recombinantly in a suitable expression system, such as a cell, by expression from encoding nucleic acid.

Following purification, the DNA-PK_{cs}-like polypeptide or protein containing the DNA-PK_{cs}-like polypeptide or the EGGDVDDLLDMI polypeptide may be used as desired, e.g. in an assay for an agent which modulates its activity, e.g.

5 binding, in raising or obtaining a specific antibody or other binding molecule, or in a therapeutic context.

Methods of determining the binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide, of identifying a 10 DNA-PK_{cs}-like polypeptide, of screening for an EGGDVDDLLDMI polypeptide able to bind to a DNA-PK_{cs}-like polypeptide, and of screening for an agent able to modulate binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide, include 15 methods in which a suitable end-point is used to assess binding.

Binding may be determined by any number of techniques known in the art, qualitative or quantitative. They include 20 techniques such as radioimmunosassay, co-immunoprecipitation, scintillation proximetry assay and ELISA methods.

Binding of a DNA-PK_{cs}-like polypeptide (e.g. DNA-PK_{cs}) to a target EGGDVDDLLDMI polypeptide (e.g. to a peptide of amino acid sequence EGGDVDDLLDMI) may be studied by labelling 25 either one with a detectable label and bringing it into contact with the other which may have been immobilised on a solid support.

Suitable detectable labels, especially for peptidyl substances include ^{35}S -methionine which may be incorporated into recombinantly produced peptides and polypeptides.

Recombinantly produced peptides and polypeptides may also be 5 expressed as a fusion protein containing an epitope which can be labelled with an antibody.

The polypeptide which is immobilized on a solid support may be immobilized using an antibody against that polypeptide 10 bound to a solid support or via other technologies which are known *per se*. A preferred *in vitro* interaction may utilise a fusion polypeptide including glutathione-S-transferase (GST). This may be immobilized on glutathione agarose beads. In an 15 *in vitro* assay format of the type described above a test modulator can be assayed by determining its ability to diminish the amount of labelled polypeptide (e.g. labelled EGGDVDDLLDMI polypeptide) which binds to the immobilized GST-fusion polypeptide (e.g. immobilised fusion polypeptide of GST and a DNA-PK_{cs}-like polypeptide). This may be determined 20 by fractionating the glutathione-agarose beads by SDS-polyacrylamide gel electrophoresis. Alternatively, the beads may be rinsed to remove unbound polypeptide and the amount of polypeptide which has bound can be determined by counting the 25 amount of label present in, for example, a suitable scintillation counter.

Binding or interaction of a DNA-PK_{cs}-like polypeptide and an EGGDVDDLLDMI polypeptide may also be determined using a two-

hybrid assay.

For example, a DNA-PK_{cs}-like polypeptide polypeptide or an EGGDVDDLLDMI polypeptide may be fused to a DNA binding domain 5 such as that of the yeast transcription factor GAL4. The GAL4 transcription factor includes two functional domains. These domains are the DNA binding domain (GAL4DBD) and the GAL4 transcriptional activation domain (GAL4TAD). By fusing the DNA-PK_{cs}-like polypeptide to one of those domains, and an 10 EGGDVDDLLDMI polypeptide to the respective counterpart, a functional GAL4 transcription factor is restored only when the DNA-PK_{cs}-like polypeptide and EGGDVDDLLDMI polypeptide interact. Thus, interaction of these polypeptides may be measured by the use of a reporter gene linked to a GAL4 DNA 15 binding site which is capable of activating transcription of said reporter gene.

This two hybrid assay format is described by Fields and Song, 1989, Nature 340; 245-246. It can be used in both mammalian 20 cells and in yeast. Other combinations of DNA binding domain and transcriptional activation domain are available in the art and may be preferred, such as the LexA DNA binding domain and the VP60 transcriptional activation domain.

25 When looking for substances which interfere with binding of (for example) a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide, a DNA-PK_{cs}-like polypeptide or a EGGDVDDLLDMI polypeptide may be employed as a fusion with (e.g.) the LexA

DNA binding domain, and the counterpart polypeptide containing the DNA-PK_{cs}-like polypeptide or EGGDVDDLLDMI polypeptide as a fusion with (e.g.) VP60. An expression cassette may be used to express a test peptide within a host 5 cell. The expression cassette may be one vector from a library of expression vectors which encode a diverse range of peptides. A reduction in reporter gene expression (e.g. in the case of β -galactosidase a weakening of the blue colour) results from the expression of a peptide which disrupts the 10 DNA-PK_{cs}-like polypeptide/EGGDVDDLLDMI polypeptide interaction, which interaction is required for transcriptional activation of the β -galactosidase gene. Where a test substance is not peptidyl and may not be expressed from encoding nucleic acid within a said third 15 expression cassette, a similar system may be employed with the test substance supplied exogenously.

An assay or screening method according to the present invention may thus take the form of an *in vivo* assay.

20

The end-point of an *in vivo* assay, that is to say the property which is determined in order to assess the binding of a DNA-PK_{cs}-like polypeptide to an EGGDVDDLLDMI polypeptide (e.g. to assess whether a test agent has an effect on the 25 binding of a DNA-PK_{cs}-like polypeptide to a EGGDVDDLLDMI polypeptide) may be a biological activity of the DNA-PK_{cs}-like polypeptide or a biological activity of the EGGDVDDLLDMI polypeptide, whichever is appropriate. As noted, those

skilled in the art well appreciate the need for and design of appropriate controls for validation of results.

As end points for in vivo assays employing human DNA-PK_{cs} or a homologue/orthologue thereof, and/or Ku, the effect on DNA repair, cell viability, cell killing, radiosensitivity, V(D)J recombination, cell cycle arrest. In yeast, P element transposition, and mating type switching may be measured.

Suitable methods are known to those skilled in the art.

10

Various further aspects and embodiments of the present invention will be apparent to those skilled in the art in view of the present disclosure. Certain aspects and embodiments of the invention will now be illustrated by way 15 of example and with reference to the figures described below.

Brief description of the Figures

Figure 1 shows a sequence alignment of Ku70 and Ku80 20 sequences. At each position in the alignment, amino acid residues with conserved identity in three or more Ku subunit proteins are shaded black, those with similarly conserved chemical or physical properties are shaded grey. The positions of the most highly conserved regions termed 25 Homology Regions 1-6 (HRs 1-6) are marked by shaded boxes above the aligned sequences. The HRs and specific C-terminal extension of Ku80 have been rendered in the same shading as used in the representations of Ku80 and Ku70 deletion mutants

shown in Figures 2 and 3. The members of the alignment are as follows: *Homo sapiens* Ku70 (HsKu70), *Gallus gallus* Ku70 (GgKu70), *Rhipicephalus appendiculatus* Ku70 (RaKu70), *Drosophila Melanogaster* Ku70 (DmKu70), *Saccharomyces cerevisiae* Ku70 (ScKu70), *Homo sapiens* Ku80 (HsKu80), *Caenorhabditis elegans* Ku80 (CeKu80) and *Saccharomyces cerevisiae* Ku80 (ScKu80).

Figure 2 shows the Ku80 deletion derivatives used to 10 investigate the regions of Ku80 able to interact with full length Ku70. The deletion derivatives were expressed as GST fusion proteins. Full length Ku80 is represented by the uppermost bar. The positions of the Homology Regions (HRs) defined in Figure 1 are indicated. For each HR, the shading 15 is the same as used in Figure 1. Deletion mutants are represented by shaded bars below this, drawn to scale with the numbers of the first and last amino acid residues of each deletion printed to the right of each bar.

20 Figure 3 shows the Ku70 deletion derivatives used to investigate the regions of Ku70 able to interact with full length Ku80. The deletion derivatives were expressed as GST fusion proteins. Full length Ku70 is represented by the uppermost bar. The positions of the Homology Regions (HRs) 25 defined in Figure 1 are indicated. For each HR, the shading is the same as used in Figure 1. Deletion mutants are represented by shaded bars below this, drawn to scale with the numbers of the first and last amino acid residues of each

deletion printed to the right of each bar.

Figure 4 is a schematic diagram of the Ku80 deletions used to define the region of Ku80 that interacts with DNA-PK_{cs}. Full length Ku80 is shown at the top with deletion mutants drawn to scale below.

Figure 5 is a schematic diagram of the C-terminus of human Ku80, amino acids 609-733, and the sequence of 5 peptides derived therefrom. These peptides, designated A-E were generated with an N-terminal biotin group followed by the sequence Ser-Gly-Ser-Gly as a linker to the Ku80 derived sequence shown.

Figure 6 is a multiple sequence alignment of the C-terminal regions of Ku80 homologues showing high sequence conservation between mammalian (*C. griseus*, *M. musculus* and *H. sapiens*) Ku80 homologues in the final 13 amino acid residues. The location of the DNA-PK_{cs} interaction motif defined within human Ku80 (the final 12 amino acid residues) is indicated.

MATERIALS AND METHODS

Yeast 2-hybrid assay

A large fragment of Ku70 (corresponding to residues 213-590) was PCR amplified and cloned into the yeast 2-hybrid "bait" vector pEG202 to generate a fusion with the DNA-binding domain Lex202 (Lex-Ku70Δ213-590), and transformed into a

yeast strain, EGY48 containing the *LEU2* gene under the control of 3 LexA operator sites integrated into the genome as well as the plasmid pSH18-34 containing the reporter gene *GAL1-Lac Z* downstream of 4 LexA operators. A HeLa cell expression library in the 2-hybrid vector pJG4-5 was 5 transformed into the yeast containing Lex-Ku70 Δ 213-590. Some of the transformation was plated onto non-selective media to calculate the transformation efficiency at $>10^7$ total transformants. Yeast transformants were subjected to a 2-fold selection procedure by first plated onto medium lacking Leucine, the picking colonies that grew up onto fresh plates containing X-gal to screen for activation of the *GAL1-LacZ* reporter. 10

15 Generation of GST-fusion proteins

Full length Ku70 and Ku80 were cloned in frame with the glutathione binding domain of human glutathione S-transferase (GST) by sub-cloning into pGEX2TKP (a derivative of pGEX2TKP by Pharmacia). Deletion mutants of Ku70 and Ku80 were cloned 20 by *Pfu* polymerase amplification and cloning into pGEX2TKP. Expression from all pGEX2TKP constructs was carried out in either DH5 α or DS941 strains of *E. coli*. Bacterial cell extracts were produced by sonication in Phosphate-buffered saline (PBS; 138 mM NaCl, 3 mM KCl, 10 mM Phosphate buffer pH 25 7.4) containing 1% Triton-X100 (Sigma) and protease inhibitor tablets (Boehringer). GST-fusion proteins were purified by passing bacterial lysates over Glutathione Agarose beads (Pharmacia) and washing extensively in PBS containing 1%

Triton-X100 and then in PBS alone. The beads/proteins were stored at -80°C prior to use.

Preparation of biotinylated bacterial cell extracts

5 Ku70 and Ku80 were cloned into pET30a (Novagen) to generate a N-terminal hexa-Histidine tag and expressed at low levels in E. coli (BL21 pLySS). The soluble protein fraction was prepared by lysing cells in buffer (50 mM phosphate pH 7.5, 150 mM NaCl, 0.1% β -mercaptoethanol, 0.1% Triton-X100, 10 Boehringer Complete Protease Inhibitors) followed by centrifugation to remove insoluble material. The soluble fraction was biotinylated by addition of biotin (Sigma) to a final concentration of 0.05 mg biotin per 1 mg extract and incubating 1 hour on ice. Unreacted biotin groups were 15 removed by addition of Glycine (0.1 M final).

GST "pulldown" assays from biotinylated extracts

Approximately 1 mg of biotinylated bacterial extract (containing His-Ku70 or His-Ku80) was diluted 2-fold in 20 buffer (50 mM Tris pH 8.0, 250 mM NaCl, 0.1% β -mercaptoethanol, 0.1% Triton X-100, Boehringer Complete Protease Inhibitors) and incubated for 3 hours at 4°C with 1-2 ug of GST-Ku80, GST-Ku70, or deletion derivative thereof, immobilised on beads. The beads were washed seven times in 25 NETN buffer (180 mM NaCl, 25 mM EDTA, 20 mM Tris pH 8.0, 0.5% NP-40), boiled in SDS gel-loading buffer and half the sample subjected to polyacrylamide gel electrophoresis. Proteins were subsequently transferred to nitrocellulose membrane.

Membranes were probed with streptavidin conjugated to horseradish peroxidase (GibcoBRL), and visualised by ECL (Pierce) and exposure to film.

5 GST "pulldown" assays from HeLa Cell Nuclear Extract
GST-Ku70 or GST-Ku80 (0.5-2.0 μ g), or deletion derivative thereof, bound to Glutathione agarose beads, was incubated with ~100 μ g (10 μ l) of HeLa Cell nuclear extract (NE) and 80 μ l of Z' buffer (25 mM HEPES pH 6.7, 12.5 mM MgCl₂, 1 mM dithiothreitol, 0.1% Nonidet P40 (BDH), 20% Glycerol)
10 containing the specified concentration of KCl (for example, 50 mM Z' contains 50 mM KCl in addition to the other buffer components). Binding reactions were incubated at 4°C for 3-4 hours and the beads recovered by centrifugation and washed 7 times in 1 ml of Buffer X (25 mM HEPES pH 6.7, 2 mM MgCl₂, 1 mM Dithiothreitol, 0.4% Nonidet P40) containing the specified concentration of KCl (for example, 120 mM Buffer X contains 120 mM KCl in addition to the other buffer components).
15 Proteins retained on the beads after washing were eluted in SDS gel-loading buffer and half of the sample subjected to polyacrylamide gel electrophoresis and silver staining.
"Pulldown" assays from purified DNA-PKcs were done as above with 400 ng of purified DNA-PKcs substituted for HeLa NE.
20

25 Peptide interaction assay

5 Peptides with the sequence Ser-Gly-Ser-Gly as a linker to 12 amino acid residues derived from the sequence of Ku80 (see Figure 5) were generated with a biotin moiety at the N-

terminus (Khiron Technologies). Each peptide (2×10^{11} moles) was incubated with either 400 ng ($\sim 10^{12}$ moles) of purified DNA-PKcs or 100 μ g of crude HeLa NE in 100 μ l (final) volume of Z' buffer (see above for buffer composition). Binding 5 reactions were incubated for 2.75 hours at 4°C, then a 100 μ l suspension of para-magnetic beads, containing 1:5 dilution of Dynabeads (Dynal) in Z' buffer, was added for a final 0.25 hour incubation. Dynabeads are conjugated with Streptavidin. The bead-protein complexes were retrieved from the 10 solution using a magnet and washed 5 times with 0.5 ml of Buffer X (see above for buffer composition). Proteins retained on the beads were eluted in SDS gel-loading buffer and subjected to polyacrylamide gel electrophoresis and silver staining

15

Purification of DNA-PKcs

The starting material for purification of DNA-PKcs was HeLa NE obtained from the Computer Cell Culture Centre, Mons, Belgium. All purification steps were performed in buffer D (20 mM HEPES pH 7.6, 2 mM MgCl₂, 0.2 mM EDTA, 10% Glycerol) with a specified concentration of KCl (for example 50 mM). Buffer D contains 50 mM KCl in addition to the other buffer components). In a representative purification, 35 ml of HeLa NE was cleared by centrifugation and diluted to a final 20 concentration of 50 mM Buffer D, and protease inhibitors added (Boehringer). This was loaded onto a 60 ml column of Q Sepharose (Pharmacia) and eluted in a linear salt gradient of 25 50-500 mM Buffer D. Under these conditions the majority of 500 mM Buffer D. Under these conditions the majority of

DNA-PKcs and Ku were separated into two fractions peaking at approximately 200 mM Buffer D, and 450 mM Buffer D respectively. The DNA-PKcs fraction was adjusted to 100 mM KCl using 0 mM Buffer D (salt concentration was measured by conductivity) and applied to a 10 ml column of heparin agarose (Sigma). The column was developed with a linear gradient of 100-200 mM Buffer D. The peak DNA-PKcs fraction from the heparin column was adjusted to 0.5 M ammonium sulphate and applied to a 2.5 ml column of phenyl Sepharose (Pharmacia) equilibrated in, typically, 300 mM Buffer D containing 0.5 M ammonium sulphate. Under these conditions, 100% of DNA-PKcs is retained by the column, whilst the majority of contaminating Ku flows through (as judged by western blotting using antibodies specific to the Ku subunits). The column was then washed extensively in 0 mM Buffer D containing 0.5 M ammonium sulphate, and subsequently in 0 mM Buffer D containing 0.4 M ammonium sulphate, which removes most of the remaining Ku, before applying a linear gradient of 0.4-0 M ammonium sulphate in 0 mM Buffer D. Peak DNA-PKcs fractions were dialysed into 50 mM Buffer D, applied to a 1 ml Mono-S FPLC column (Pharmacia), and eluted in a linear gradient of 50-500 mM Buffer D. Peak fractions were dialysed into 50 mM Buffer D and stored at -80°C prior to use. The resulting protein preparation contained a single abundant polypeptide corresponding to DNA-PKcs as judged by Coomassie staining, silver staining and western blotting, with a ladder of smaller DNA-PKcs breakdown products. Western blotting of a variety of HeLa cell extracts using antibodies specific to

DNA-PKcs reveals an essentially identical pattern of ladder, hence we believe that the majority of these breakdown products are present at the start of the DNA-PKcs preparation. To assess the levels of Ku contamination in the 5 DNA-PKcs preparation we carried out quantitative western blotting using Ku-specific antibodies. In these western blots, an essentially homogeneous preparation of Ku was titrated against known amounts of our DNA-PKcs preparations. The resulting signal suggested a typical contamination level 10 of 1 molecule of Ku to 840 molecules of DNA-PKcs.

RESULTS

A C-terminal region of Ku80 interacts with Ku70 in the yeast 15 2-hybrid assay

The inventors used a yeast 2-hybrid assay to screen a human HeLa cell cDNA library for proteins that interact with human Ku70. A large fragment of Ku70 (spanning residues 434 to 20 733), containing two putative leucine zipper motifs, was used as the "bait". A two-fold selection procedure was used (see Materials and Methods). From the first selection 530 positive colonies were picked and plated onto X-gal plates to assay for β -galactosidase activity. 55 positive clones were identified and the library plasmids contained within these 25 clones were rescued into *E. coli*. Notably, bacterial colony hybridisation to a radiolabelled Ku80 DNA probe revealed that 44% gave a strong positive signal, and when these clones were sequenced, all were found to contain the Ku80 cDNA. The

retrieved plasmids comprised eight distinct Ku80 cDNA clones: all contained the poly(A) tail of the Ku80 cDNA but differed in their 5' ends. The different 5' termini were highly clustered, all falling to within 27 nucleotides of each other. The smallest clone encodes the C-terminal 293 amino acid residues of Ku80, defining this region as sufficient for interaction with Ku70 in this assay.

10 Analysis of interactions between bacterially-expressed Ku70 and Ku80 deletion derivatives

In a complementary approach to the yeast 2-hybrid technique, the inventors used a "pull-down" assay to detect interactions between recombinant deletion derivatives of the two Ku subunits. To first establish the utility of this approach, 15 they assayed for interaction between full-length Ku70 and Ku80 in this system. Both Ku subunits were sub-cloned into bacterial expression vectors designed to express them either as a GST-fusion or with an N-terminal hexa-histidine tag (see Materials and Methods). GST-Ku70 and GST-Ku80 were expressed 20 in *E. coli* and purified on glutathione-agarose beads. His-Tagged Ku subunits were also expressed, and the crude unfractionated bacterial cell lysates containing either His-tagged Ku70 or His-tagged Ku80 were biotinylated at low levels (see Materials and Methods). Interaction assays were 25 then conducted by incubating either GST-Ku80 beads with bacterial lysate containing His-Ku70, or GST-Ku70 beads with His-Ku80 lysate. The beads were next washed extensively, and then any proteins that remained bound to them were eluted in

SDS-PAGE buffer, resolved on a polyacrylamide gel and western blotted. Interacting proteins were detected by probing these western blots with streptavidin that was conjugated to horse radish peroxidase (HRP). Control binding reactions were 5 carried out using GST-alone bound beads (at a five fold molar excess over GST-Ku70/Ku80) and bacterial lysates containing either His-tagged Ku80 or His-tagged Ku70.

Of the large number of biotinylated proteins present in crude 10 extracts, only a single species is retained efficiently on the beads in pull-downs with either GST-Ku70 or GST-Ku80. The identity of the interacting proteins as Ku80 and Ku70, respectively, was confirmed by quantitative western blotting with anti-Ku subunit antisera. In the control binding 15 reactions using GST alone, neither Ku subunit was retained from the bacterial lysate. Although the above results suggest that the assay is detecting tight and highly specific protein-protein interactions between the two Ku subunits, another possible explanation was that the interaction might 20 be being mediated via DNA that could have been present in the protein preparations. To address this issue, the inventors performed the above binding assays in the presence of high amounts of ethidium bromide (EtBr), which disrupts complexes between DNA and DNA-binding proteins (Lai and Herr, 1992). 25 Importantly, the presence of EtBr has no effect on the Ku subunit interactions in the assay, demonstrating that the interaction being measured is direct.

Sequence analysis of Ku70 and Ku80 indicates that they are structurally and evolutionarily related

The inventors analysed the human Ku70 and Ku80 protein sequences for candidate motifs involved in their interaction.

5 They noted that the two subunits share regions of sequence similarity (a subset of which have been detected recently (Dynan and Yoo, 1998)) and that these similarities are shared by all the known homologues of Ku70 and Ku80. A Multiple Sequence Alignment of selected Ku70 and Ku80 sequences was

10 generated using the program "pileup" from the sequence analysis "Wisconsin Package, version 8.1" (Program Manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group: Gap creation 3.00; Gap extension 0.10). The alignment is shown in Figure 1: sequences from diverse

15 branches of life were selected in order to give an indication of the most highly conserved regions, hence mouse or hamster homologues of Ku70 and Ku80 have not been included as these are almost identical to the human sequences.

20 Significantly, all of the Ku sequences align along their entire lengths, without necessitating the introduction of large gaps (the only notable exception to this occurs where the C-termini of the *Caenorhabditis elegans* and Human Ku80 sequences extend beyond the other sequences; see below).

25 Within the aligned region, the inventors established six segments of significant similarity between all of the Ku sequences, which they have designated Homology Regions 1-6 (HRS 1-6). These occur in the same order and with conserved

spacing in each Ku sequence.

The features of the Ku70 and Ku80 sequences noted above suggest that these molecules fold into similar structures and 5 may have evolved from a common ancestor. The HRs may correspond to conserved functional domains or structural features of the Ku70 and Ku80 polypeptides. Such regions could play key roles in specifying correct protein folding or mediating protein-protein and/or protein/nucleic acid 10 interactions. Secondary structure predictions assign a high probability of α -helix formation within HRs 3, 5 and 6, and of β -strand formation in HR1. The short regions that separate HRs 2-6 are of variable length in the different 15 proteins. These regions may correspond to inter-domain linkers or variable loops on the protein surface.

Interestingly, within the region separating HR1 and HR2, there is very little similarity between the Ku70 and Ku80 sequences but there are several regions of sequence 20 similarity exclusive to either the Ku70 or the Ku80 sequence families. Such regions could, therefore, specify Ku subunit-specific functions, although it is also possible that they share significant structural similarity with the other Ku subunit, but this is not easily detectable at the primary 25 amino acid sequence level.

25

A second region exhibiting differences between the Ku70 and Ku80 sequences corresponds to the proteins' extreme C-termini. Thus, *C. elegans* and human Ku80 sequences are

considerably longer than other sequences in the alignment, and contain C-terminal extensions of 81 and 108 amino acid residues, respectively. The larger size of Ku80 compared to Ku70 is, therefore, largely due to the presence of a larger

5 C-terminal domain (or an extra domain at the C-terminus).

Throughout the present disclosure, this region is referred to as the Ku80 specific C-terminal extension (Ku80-CTE). The Ku70 sequences also share an exclusive extreme C-terminal sequence motif. The region that the inventors have

10 identified as interacting with Ku70 in the 2-hybrid assay

includes HR5 and HR6, making these attractive candidates for pseudo-homodimerisation interaction motifs.

Analysis of regions of Ku80 required for heterodimerisation

15 The inventors tested a range of Ku80 deletion derivatives expressed as GST-fusions for their ability to interact with full-length Ku70. These derivatives and their relationships to the Ku80 homology regions are depicted in Figure 2. The Ku80 derivatives were tested for an ability to affinity-

20 purify full-length Ku70 from a crude bacterial extract, as

described above. Briefly, GST-Ku80 derivatives were bound to glutathione agarose beads and incubated with biotinylated bacterial cell extracts containing full length Ku70.

Proteins retained on the beads after seven washes in NETN

25 buffer (20mM Tris pH8.0, 150mM KCL, 1mM EDTA, 0.5% NP-40)

were analysed by SDS-PAGE, western blotting and probing with HRP-conjugated streptavidin.

These binding studies revealed that Ku80 Δ N1, Ku80 Δ N2 and Ku80 Δ N3 interact with Ku70, but Ku80 Δ N3 interacts more strongly than the other two, larger, derivatives. These observations provide a potential explanation for why Ku80 derivatives longer than Ku80 Δ N3 were not isolated in the 2-hybrid screen, and raise the possibility that sequences in HR3 and HR4 of Ku80 modulate the accessibility of the Ku70 interaction domain.

Notably, further N-terminal deletions of Ku80, represented by Ku80 Δ N4 and Ku80 Δ N5, reduce the binding of Ku70 to undetectable levels. Thus, residues in Ku80 HR5 are necessary for interaction with Ku70 in these assays. C-terminal Ku80 truncation derivatives were also tested. These analyses revealed that Ku80-CTE, the C-terminal region of human Ku80 that has no clear homology to Ku70, is not required for the Ku70 interaction. In contrast, further C-terminal deletion, represented by Ku80 Δ N1C2, virtually eliminates Ku70 binding. Finally, a Ku80 derivative, Ku80 Δ N3C1, which comprises only amino acid residues 439-592, is capable of binding Ku70 with high efficiency and specificity. These data reveal that the region of Ku80 comprising HR5 and HR6, together with an additional region between HR6 and the Ku80-CTE, mediates highly specific binding to Ku70.

Analysis of regions of Ku70 required for heterodimerisation.
The inventors generated a series of GST-Ku70 deletion mutants

and tested their ability to affinity-purify Ku80 from a biotinylated crude bacterial cell lysate containing full-length Ku80. The Ku70 derivatives and their relationships to the Ku homology boxes are depicted in Figure 3. Once again, 5 GST-Ku70 derivatives were bound to glutathione agarose beads and incubated with biotinylated bacterial cell extracts containing full length Ku80. Proteins retained on the beads after seven washes in NETN buffer (20mM Tris pH8.0, 150mM KCL, 1mM EDTA, 0.5% NP-40) were analysed by SDS-PAGE, western 10 blotting and probing with HRP-conjugated streptavidin.

Full-length Ku70 binds to Ku80 efficiently, but does not interact with any of the other proteins present in crude bacterial extract. Whereas the N-terminal Ku70 region 15 (Ku70 Δ C5) that contains HR1 and HR2 displays very little binding to Ku80, the C-terminal region (Ku70 Δ N1) containing HRs 3-6 binds Ku80 with high efficiency and specificity. Subdivision of Ku70 Δ N1 allowed the interaction region of Ku70 to be further defined. Thus, whereas deletion of HR3 and 20 HR4, exemplified by Ku70 Δ N2 and Ku70 Δ N3, has no discernible effect on Ku80 binding, further N-terminal deletion derivatives lacking HR5 or HRs 5 and 6 (Ku70 Δ N4 and Ku70 Δ N5, respectively) are totally abrogated for the Ku80 interaction. 25 C-terminal deletion of Ku70 Δ N1 revealed that truncation of the final 31 amino acid residues has no discernible effect on Ku80 binding. This C-terminal region corresponds to a sequence motif that is apparently not present in *S.*

cerevisiae Ku70 but exists in all other Ku70 homologues (Figure 1), an observation which is consistent with the finding that this region does not function in Ku subunit interaction, and suggests a function for mammalian and 5 arthropod Ku70 that does not exist in yeast. Any further deletion of the Ku70 C-terminus, however, leads to dramatic reduction in Ku80 binding.

Taken together, the above data indicate that a region of Ku70 10 containing HR5 and HR6, along with a third region of low Ku70/80 homology immediately C-terminal to HR6, is capable of binding to Ku80 with high specificity. It is striking that the interaction regions that the inventors have defined for both human Ku70 and Ku80 share significant similarities in 15 sequence and position within the Ku polypeptides. The present disclosure therefore provides support for a model in which the Ku subunits interact by a pseudo-homodimerisation mechanism.

20 The present inventors have investigated Ku subunit interactions by generating untagged deletion derivatives of Ku70 and Ku80 using *in vitro* transcription and translation in rabbit reticulocyte lysate, followed by co-immunoprecipitation using antibodies specific for Ku70. 25 These studies indicate that HR5 and HR6 are involved in the interaction, but show that a further 96 N-terminal amino acid residues, comprising HR3 and HR4, are also required in this system.

Ku80 interacts directly with DNA-PKcs

The inventors performed pull-down assays from HeLa cell nuclear extract using GST-Ku70 or GST-Ku80 bound to glutathione agarose beads. Thus, GST-Ku70 or GST-Ku80 5 protein was purified on glutathione agarose beads as described above, and the beads incubated with crude HeLa cell nuclear extract and washed extensively in buffer X containing 150 mM KCl. The bound proteins were then eluted in the presence of SDS, subjected to SDS-PAGE on a 6% polyacrylamide 10 gel and detected by silver-staining. A control binding reaction was carried out using beads containing only GST.

Of the many proteins in the crude extract, DNA-PKcs is the only protein that is retrieved by affinity purification with 15 either GST-Ku70 or GST-Ku80 beads (the identity of the retrieved species as DNA-PKcs was verified by quantitative western blot analyses). Indeed, other than proteins derived from the GST-Ku70 or GST-Ku80 preparations, DNA-PKcs is essentially the only polypeptide in the retrieved fractions. 20 In contrast, DNA-PKcs is not recovered using beads containing GST alone.

The inventors investigated whether the observed interactions between DNA-PKcs and the individual Ku subunits were indirect 25 and were being mediated by DNA and/or the Ku heterodimer present in the HeLa nuclear extract. EtBr was used in a final concentration of 0.2mg/ml in order to address the potential involvement of DNA (EtBr disrupts protein-DNA

interactions). Notably, the interaction between Ku80 and DNA-PK_{cs} is unaffected by EtBr, suggesting that it is either direct or is mediated via a protein-protein interaction that does not require DNA. In contrast, the interaction between 5 DNA-PK_{cs} and Ku70 is essentially abolished by EtBr treatment, indicating that the interaction between Ku70 and DNA-PK_{cs} is dependent upon protein-DNA interactions.

The inventors carried out binding reactions using an 10 essentially homogeneous preparation of DNA-PK_{cs} (this preparation was shown to contain Ku contamination at a level of around 1 in 840 DNA-PK_{cs} molecules; see Materials and Methods). GST-Ku70 or GST-Ku80 containing glutathione agarose beads were incubated with 0.4 μ g of purified DNA-PK_{cs} 15 under conditions of 120 mM KCl. The beads were then washed and any proteins retained on the beads were detected by silver-staining.

To investigate the effect of DNA upon these interactions 20 binding reactions with purified DNA-PK_{cs} were carried out in either in the presence of 20ng/ μ l sonicated calf thymus DNA (which contains dsDNA ends and other discontinuities in the DNA double-helix and was shown to be highly effective at activating DNA-PK catalytic activity) or in the presence of 25 0.2mg/ml EtBr. A control binding reaction using beads containing GST alone was also carried out.

Significantly, these studies revealed that GST-Ku80 interacts

efficiently with purified DNA-PKcs and that this interaction is not affected by EtBr or by sheared genomic DNA.

Furthermore, the inventors were unable to detect any endogenous Ku contamination in these Ku80 pull-downs using

5 either polyclonal rabbit sera or monoclonal antibodies raised against the Ku subunits. Taken together, the data therefore reveal that there is a specific and direct interaction between Ku80 and DNA-PKcs, and eliminate the possibility that another protein species might be contributing to the 10 interactions observed using the HeLa nuclear extract.

In contrast to the above results, very little interaction between Ku70 and purified DNA-PKcs was detected.

Furthermore, no stimulation of binding was observed in the 15 presence of DNA, arguing that the binding of GST-Ku70 to DNA-PKcs in the crude nuclear extract cannot be explained by DNA directly bridging between Ku70 and DNA-PKcs, nor by it inducing a conformational change in Ku70 or DNA-PKcs.

20 The inventors conducted assays with homogeneous DNA-PKcs and either GST-Ku70 or GST-Ku80 under conditions ranging from 50 mM to 150 mM KCl. These studies revealed that, at lower but not at higher ionic strengths, there is an interaction between Ku70 and DNA-PKcs. The interaction between DNA-PKcs 25 and the Ku80 subunit, however, is much more salt stable and is still easily detectable at 150 mM KCl. Importantly, no significant interaction between GST and DNA-PKcs was observed at any salt concentration used, and the addition of DNA or

EtBr had no affect on the binding of GST-Ku70 or GST-Ku80 under any of the conditions employed. The inventors conclude that there are contacts between each of the two Ku subunits and DNA-PKcs, but the contribution made by the Ku80 subunit 5 is of greater strength and likely significance, given that it is the only interaction that can be detected at physiological salt concentrations.

The Ku80 C-terminal region interacts with DNA-PKcs

10 The inventors performed interaction assays with purified DNA-PKcs (to avoid the possible contribution of endogenous Ku or other protein species in nuclear extract) and a battery of Ku80 deletion derivatives that had been expressed as GST-fusion proteins and bound to the glutathione-agarose beads.

15 Binding reactions and subsequent washes were carried out in NETN buffer (20mM Hepes pH 7.5, 2mM MgCl₂, 0.4% NP-40) containing 120 mM KCl to assay for the stronger, likely more physiological, binding observed between Ku80 and DNA-PKcs.

20 Bound proteins were resolved on a 7% polyacrylamide gel and were detected by silver-staining. The Ku80 deletion derivatives used are shown in Figure 4.

It was found that sequential deletions from the N-terminus of Ku80 (Ku80 Δ N1-N6) have little or no effect on DNA-PKcs 25 binding. Significantly, the smallest construct in this series, Ku80 Δ N6, comprises amino acid residues 595-732 of Ku80 and corresponds to the C-terminal region of human Ku80, the Ku80-CTE, which has no homology to Ku70 sequences (see

Figure 1).

The expression construct encoding Ku80 Δ N8 was designed to express only 22 amino acid residues corresponding to the 5 extreme C-terminus of Ku80. Ku80 Δ N7 and Ku80 Δ N8 are still able to mediate an effective interaction with DNA-PKcs.

The inventors generated Ku80 derivatives representing a series of deletions from the C-terminus. These derivatives 10 are shown in Figure 4: they lack 0, 28, 70, 140, 202 or 267 amino acid residues from the C-terminus. Strikingly, removal of only 28 amino acid residues renders the resulting Ku80 derivative protein unable to bind DNA-PKcs.

15 Taken together, the data reveal that the interaction between human Ku80 and DNA-PKcs occurs through the Ku80-CTE region, and that the final 28 amino acid residues of Ku80 are necessary and sufficient for this interaction.

20 A peptide from the extreme C-terminus of Ku80 is sufficient to mediate highly specific and efficient interactions with DNA-PKcs

The inventors synthesised a series of peptides (Peptides A-E) derived in sequence from the extreme C-terminus of Ku80 25 (residues 609 to 733: see Figure 5 and Materials and Methods). These peptides each contained 12 amino acid residues derived from overlapping sequence segments of Ku80, and have an N-terminal biotin moiety separated from the Ku80

derived sequence by the linker sequence Ser-Gly-Ser-Gly. Each peptide (2×10^{-11} moles) was incubated with homogenous preparation DNA-PKcs (1×10^{-12} moles) in either Z' buffer containing 50mM KCl or Z' buffer containing 100mM KCl. The 5 inventors then added paramagnetic beads conjugated with streptavidin to capture the peptide via the biotin-streptavidin interaction. The beads were recovered using a magnet, washed extensively in buffer containing the same concentration of KCl as the incubation buffer X (25mM HEPES 10 pH 7.5, 50mM/100mM KCl, 2mM MgCl₂, 0.4% NP-40), and the bound protein eluted in SDS-PAGE sample buffer, resolved on a 6% polyacrylamide gel and detected using silver-staining.

Peptide E, comprising the final 12 amino acid residues at the 15 C-terminus of Ku80, interacts directly with DNA-PKcs at both 50 mM and 100 mM KCl. In contrast, none of the other peptides (Peptides A-D) exhibit significant DNA-PKcs binding.

The ability of each of the five peptides to retrieve proteins 20 from a crude unfractionated HeLa cell nuclear extract was determined. Each peptide was incubated with 100mg of HeLa nuclear extract in 50mM Z' buffer, followed by capture of the peptide onto streptavidin-conjugated magnetic beads and subsequent washing in buffer X containing 50mM KCl. Any 25 proteins remaining bound to the peptide/beads was eluted in SDS-PAGE buffer, resolved on a 6% polyacrylamide gel and detected with silver-staining. A control binding reaction was carried out with beads alone.

Strikingly, Peptide E affinity purified predominantly a single protein species of high molecular weight. This retrieved protein was confirmed to be DNA-PKcs by quantitative western blotting. Furthermore, western blotting 5 revealed that the faint bands seen at around 150 kDa and 105 kDa correspond to breakdown products of DNA-PKcs. Additional studies have shown that Peptide E is able to deplete 60-70% of DNA-PKcs from HeLa nuclear extract, but that some DNA-PKcs always remains unbound, and that both DNA-PKcs affinity 10 purified on Peptide E and that remaining in the extract display DNA-end stimulated kinase activity upon addition of purified Ku. In contrast, Peptides A-D and a variety of control peptides all fail to specifically retrieve any proteins from nuclear extract.

15

Taken together, these results demonstrate that the extreme C-terminal 12 amino acid residues of Ku80 are sufficient to mediate a highly specific interaction with DNA-PKcs.

20 *DISCUSSION*

The present disclosure shows that Ku70 and Ku80 share considerable amino acid sequence similarity, suggesting that these two proteins have similar structures. Moreover, the 25 inventors have identified a region in Ku70 (amino acid residues 449-578) and one in Ku80 (residues 439-592) that mediate contacts within the Ku heterodimer. Strikingly, the regions of Ku70 and Ku80 that interact with one another share

sequence homologies (HR5 and HR6; see Figure 1) and are therefore predicted to be analogous structures. These data strongly suggest that the Ku subunits interact with one another through a pseudo-homodimerisation mechanism.

5

There is good agreement between the results presented herein and those of other researchers regarding the N-terminal boundary of the Ku subunit interaction domains (Osipovich et al., 1997; Wang et al., 1998a; Wu and Lieber, 1996).

10 Furthermore, the C-terminal boundaries for the interaction domains have been located previously to amino acid residue 531 of human Ku80 (Osipovich et al., 1997) and residue 519 of human Ku70 (Cary et al., 1998) - in each case, these lie directly C-terminal to HR6. The data of the present inventors, however, indicate that other residues C-terminal to HR6 are also required for effective Ku70-Ku80 binding. Importantly, the results of previous point-mutational and functional analyses of the Ku subunits (Jin, S.F., et al. 15 (1997) *Embo Journal* 16, 6874-6885) retrospectively provide significant support for the homologies between Ku70 and Ku80 being of functional significance (Jin and Weaver, 1997). Thus, mutation of residues that are highly conserved in our 20 sequence analyses tend to abrogate Ku function, whereas mutation of those that are not conserved largely have little 25 or no effect.

By generating untagged deletion derivatives of Ku70 and Ku80 using *in vitro* transcription and translation in rabbit

reticulocyte lysate, followed by co-immunoprecipitation using antibodies specific for Ku70 the present inventors have confirmed that HR5 and HR6 are involved in effective interaction of the Ku subunits. These studies further 5 revealed that a further 96 N-terminal amino acid residues, comprising HR3 and HR4, are required in this system. In this connection, it is noteworthy that Cary et. al. have reported that the Ku80 dimerisation domain maps to a central region of the protein (aa241-555: Cary et al., 1998), and that this 10 region contains HR3 and HR4, in addition to HR5 and HR6. From Figure 1 it is clear that HR3 and HR4 comprise the region that is most highly conserved between Ku70 and Ku80. This raises the possibility that HR3 and HR4 form a structural core within the Ku70-Ku80 heterodimer which is 15 required for effective interactions between untagged derivatives of Ku70 or Ku80, but not when such derivatives are expressed fused to other structural motifs (GST or certain 2-hybrid fusion domains). 20 Previous studies have provided little information about the interaction of Ku with DNA-PKcs. The data presented herein demonstrate that both Ku80 and Ku70 are independently able to interact with DNA-PKcs, although the Ku80 interaction is much stronger at physiologically-relevant salt concentrations. 25 Furthermore, the inventors have shown that the major DNA-PKcs binding domain of Ku80 is located at the extreme C-terminus of the molecule, and that the final C-terminal 12 amino acid residues of Ku80 are sufficient to mediate this highly

specific interaction. This interaction is however weaker than that observed with a slightly larger C-terminal Ku80 fragment, suggesting that further residues in the Ku80-CTE might be required to mediate full-strength DNA-PKcs binding.

5

There are several lines of evidence to support the relevance of the Ku80-DNA-PKcs interaction disclosed herein *in vivo*. First, binding occurs under ionic conditions that are close to physiological. Second, this interaction is much stronger than the binding of DNA-PKcs to any other region of either Ku subunit, and the inventors have been unable to detect binding of DNA-PKcs to other control proteins or peptides. Third, the interaction between DNA-PKcs the C-terminal region of Ku80, either expressed as a GST fusion or presented as a biotinylated peptide, is highly selective, and can be used to affinity-purify DNA-PKcs to virtual homogeneity in one step from crude unfractionated human nuclear extract.

Perhaps surprisingly, GST-Ku80 and GST-Ku80 derivatives containing the C-terminal interaction domain identified by the present inventors bind to DNA-PKcs independently of dsDNA ends. This contrasts with the interaction between the native Ku heterodimer and DNA-PKcs, which only occurs in the presence of DNA (Gottlieb and Jackson, 1993; Suwa et al., 1994). Although other possibilities exist, one model to explain this difference is that the DNA-PKcs interaction region of Ku80 is masked in the native Ku heterodimer and only becomes exposed when a conformational change takes place

upon DNA binding.

An exciting aspect arising from the work present herein is that DNA-PKcs-binding peptides, e.g. Peptide E (EGGDVDDLLDMI), can function as specific inhibitors of DNA-PK by blocking the interaction between the Ku heterodimer and DNA-PKcs. There are currently a number of compounds, such as wortmannin and LY294002, that specifically inhibit members of the PIKL protein kinase and PI3-kinase families but until now 10 there were no inhibitors specific for DNA-PKcs.

The sequence analysis shows that Ku80 and Ku70 differ dramatically at their C-termini. Ku80 homologues contain a C-terminal extension (CTE) region and Ku70 homologues appear 15 to have a small Ku70-specific C-terminal region (See Figure 1). These Ku subunit-specific regions appear not to participate in homodimerisation or DNA-binding and are presumably either modifications of the ancestral C-terminal domain, or additional protein domains. Consistent with this, 20 the location of the DNA-PKcs interaction site in the Ku80-CTE suggests that the C-terminal domains form structures that are exposed on the surface of the Ku heterodimer, and hence this region of the protein can evolve an additional functionality to the Ku heterodimer without deleteriously affecting the 25 DNA-binding function.

Considering that the final 12 residues of Ku80 are making contacts with DNA-PKcs, the large size of the Ku80-CTE in

human Ku80 may allow Ku to overcome the steric hindrance which may inherently be a problem in interacting with a protein the size of DNA-PKcs. A second explanation may involve the modulation of DNA-PKcs binding by conformational

5 change of the Ku80-CTE upon DNA-binding (see above).

Alternatively the Ku80CTE (and also the Ku70-specific C-terminal region) may be the site of other protein-protein interactions. For example, it is noteworthy that, although lacking the DNA-PKcs interaction motif, *C. elegans* Ku80 does possess a CTE, and this has several regions of homology with the Ku80 CTEs of mammals. It is tempting to speculate that these regions contact other proteins involved in DNA DSB repair or DNA damage signalling, or play roles in additional functions that have been ascribed to Ku, such as regulating telomeric functions and controlling chromatin structure

(Featherstone and Jackson, 1999; Smith and Jackson, 1999).

Analyses of the Ku70 and Ku80 CTEs, and the other homology regions of these two proteins that we have identified, are

likely to be instrumental in furthering the understanding of

20 the functions of Ku in these important processes.

The Ku heterodimer exists in cells from many organisms, including mammals, flies, nematode worms and yeast. However, no clear Ku homologues exist in the fully sequenced genomes of various Bacteria and Archaea. This suggests that Ku evolved subsequently to the divergence of the eukaryotic and prokaryotic lineages, and clear sequence similarities between the Ku70 and Ku80 proteins strongly suggest they arose by a

gene duplication and were already present in the last common ancestor of *S cerevisiae* and metazoa (it seems likely, therefore, that Ku homologues will also exist in metazoan plants).

5

For DNA-PKcs, however, the situation is somewhat different; genes for DNA-PKcs orthologues have been found in human, mouse, hamster, horse and *Xenopus laevis* (Dynan and Yoo, 1998) but are not evident in *S. cerevisiae*, nor in the 10 virtually complete genome sequence of *C. elegans*.

Significantly, the presence or absence of DNA-PKcs is mirrored by the state of the C-terminus of Ku80 in these organisms. Thus, the final 12 amino acid residues of human Ku80 that we have identified as interacting with DNA-PKcs are 15 strikingly conserved in mouse and hamster Ku80 and these residues are absent from the *C. elegans* and *S. cerevisiae* Ku80 (or *S. cerevisiae* Ku70) homologues.

Furthermore the inventors have identified an open reading 20 frame in the DNA database encoding a putative homologue of Ku80 from *Drosophila* that also lacks a DNA-PKcs interaction motif, highly suggestive that DNA-PK is absent from *Drosophila*. Hence, the first known occurrence of DNA-PKcs in evolution is in the vertebrates.

25

The above observations suggest that the DNA-PK enzyme has evolved to function in DNA DSB repair by hijacking an existing component of the Non-homologous end-joining (NHEJ)

machinery, namely Ku, and hence adding to, or altering the functions that Ku can play in DNA repair. The function of Ku in NHEJ is conserved from yeast to humans but is only one of the two known pathways for the repair of DNA DSBs, the second 5 mechanism being homologous recombination with an undamaged DNA strand.

All organisms studied appear to have the machinery for both NHEJ and homologous recombination (Critchlow, S.E., et al 10 (1998) TIBS 23, 394-398, and references cited therein) however there is a distinct difference in the preference for these two DNA repair alternatives between yeast and humans. Yeast predominantly utilise homologous recombination whereas mammalian cells favour NHEJ. Although the high efficiency of 15 NHEJ in *Xenopus* is well documented, the relative importance of NHEJ and recombination in this and other organisms is less well studied than in yeast or man. There is at least a correlation between the existance of DNA-PKcs in mammalian cells and a switch to NHEJ as the major pathway of DNA DSB 20 repair.

Although it is true that NHEJ and homologous recombination are both used for the repair of randomly generated DNA damage, the process of V(D)J recombination, which involves 25 creation of dsDNA breaks at specific sites in the immunoglobulin gene loci, only occurs through a mechanism of direct end joining. DNA-PK is required for V(D)J recombination and evolution of DNA-PK may reflect the

requirement for the specific direct end-joining activity required in the generation of antibody diversity. Perhaps the specialisation of DNA-repair machinery of vertebrate cells for this specific form of direct end joining resulted 5 in a general reduction of homologous recombination mechanisms in favour of NHEJ, or perhaps V(D)J recombination was able to develop due to the ability of cells to efficiently perform NHEJ. It has been suggested that the mechanism of V(D)J recombination has arisen after the divergence of the jawed 10 and jawless vertebrates due to the insertion of a transposable element, that encoded the ancestral RAG1/RAG2 genes, into an ancestral receptor gene resembling the immunoglobulin and T-cell receptor genes found in present day jawed vertebrates. Hence it would be interesting to 15 determine the time of appearance of the DNA-PKcs gene with respect to this event.

REFERENCES

Baumann and West (1998) *Proceedings Of the National Academy Of Sciences Of the United States Of America*, 95, 14066-14070.

5 Blier et al., (1993) *Journal Of Biological Chemistry*, 268, 7594-7601.

Bogue et al., (1998) *Proceedings Of the National Academy Of Sciences Of the United States Of America*, 95, 15559-15564.

Boulton and Jackson (1996a) *Nucleic Acids Research*, 24, 4639-10 4648.

Boulton and Jackson (1996b) *Embo Journal*, 15, 5093-5103.

Boulton and Jackson (1998) *Embo Journal*, 17, 1819-1828.

Cary et al., (1998) *Nucleic Acids Research*, 26, 974-979.

Cary et al., (1997) *Proceedings Of the National Academy Of Sciences Of the United States Of America*, 94, 4267-4272.

15 Chan and Leesmiller (1996) *Journal Of Biological Chemistry*, 271, 8936-8941.

Chou et al., (1992) *Journal Of Experimental Medicine*, 175, 1677-1684.

20 Critchlow and Jackson (1998) *Trends In Biochemical Sciences*, 23, 394-398.

Devries et al., (1989) *Journal Of Molecular Biology*, 208, 65-78.

Dvir et al., (1992) *Proceedings Of the National Academy Of Sciences Of the United States Of America*, 89, 11920-11924.

25 Dynan and Yoo (1998) *Nucleic Acids Research*, 26, 1551-1559.

Falzon et al., (1993) *Journal Of Biological Chemistry*, 268, 10546-10552.

Featherstone and Jackson (1999) *Mutation Research - DNA repair, in press.*

Finnie et al., (1995) *Proceedings Of the National Academy Of Sciences Of the United States Of America, 92*, 320-324.

5 Giffin et al., (1996) *Nature, 380*, 265-268.

Gottlieb and Jackson (1993) *Cell, 72*, 131-142.

Gu et al., (1996) *Journal Of Biological Chemistry, 271*, 19660-19663.

Gu et al., (1998) *Biochemistry, 37*, 9827-9835.

10 Gu et al., (1997) *Proceedings Of the National Academy Of Sciences Of the United States Of America, 94*, 8076-8081.

Hammarsten and Chu (1998) *Proceedings Of the National Academy Of Sciences Of the United States Of America, 95*, 525-530.

Hartley et al., (1995) *Cell, 82*, 849-856.

15 Jin and Weaver (1997) *Embo Journal, 16*, 6874-6885.

Kirchgessner et al., (1995) *Science, 267*, 1178-1183.

Koike et al., (1998) *Biochemical and Biophysical Research Communications, 252*, 679-685.

Lai and Herr (1992) *Proceedings Of the National Academy Of Sciences Of the United States Of America, 89*, 6958-6962.

20 Leesmiller et al., (1995) *Science, 267*, 1183-1185.

Mimori and Hardin (1986) *Journal Of Biological Chemistry, 261*, 375-379.

Oettinger et al., (1990) *Science, 248*, 1517-1523.

25 Osipovich et al., (1997) *Journal Of Biological Chemistry, 272*, 27259-27265.

Paillard and Strauss (1991) *Nucleic Acids Research, 19*, 5619-5624.

Porter *et al.*, (1996) *Nucleic Acids Research*, **24**, 582-585.

Schatz *et al.*, (1989) *Cell*, **59**, 1035-1048.

Siede *et al.*, (1996) *Genetics*, **142**, 91-102.

Singleton *et al.*, (1999) *Molecular and Cellular Biology*, **in press**.

Sipley *et al.*, (1995) *Proceedings Of the National Academy Of Sciences Of the United States Of America*, **92**, 7515-7519.

Smith and Jackson (1999) *Genes and Development*, **submitted**.

Suwa *et al.*, (1994) *Proceedings Of the National Academy Of Sciences Of the United States Of America*, **91**, 6904-6908.

Taccioli *et al.*, (1994) *Science*, **265**, 1442-1445.

Wang *et al.*, (1998a) *Journal Of Biological Chemistry*, **273**, 842-848.

Wang *et al.*, (1998b) *Journal Of Biological Chemistry*, **273**, 31068-31074.

Wang *et al.*, (1994) *Febs Letters*, **351**, 219-224.

Wu and Lieber (1996) *Molecular and Cellular Biology*, **16**, 5186-5193.

Yaneva *et al.*, (1997) *Embo Journal*, **16**, 5098-5112.

Zhu *et al.*, (1996) *Cell*, **86**, 379-389.

CLAIMS:

1. A screening method for identifying or obtaining a binding partner of Ku70 or Ku80, said method comprising;

(a) bringing a test binding partner into contact with one or 5 more of;

i) a peptide consisting of an amino acid sequence which shares at least 70% homology with an amino acid sequence of a homology region of Ku70 or Ku80 as shown in Figure 1,

10 ii) a peptide consisting of an amino acid sequence which shares at least 70% homology with the amino acid sequence EGGDVDDLLDMI,

(b) determining binding of the test binding partner to one or more of said peptides.

15

2. A method of screening for an agent which modulates interaction between Ku70 and Ku80 and a binding partner, comprising,

(a) bringing a binding partner of Ku70 or Ku80 into contact 20 with one or more of:

i) a peptide consisting of an amino acid sequence which shares at least 70% homology with an amino acid sequence of a homology region of Ku70 or Ku80 as shown in Figure 1;

25 ii) a peptide consisting of an amino acid sequence which shares at least 70% homology with the amino acid sequence EGGDVDDLLDMI,

in the presence of a test agent;

(b) determining binding of said one or more peptides to the binding partner.

3. A screening method according to claim 1 or claim 2
5 wherein said binding partner of Ku70 or Ku80 is brought into contact with one or more of:

i) a peptide consisting of an amino acid sequence of a homology region of Ku70 or Ku80 as shown in Figure 1;
ii) a peptide consisting of the amino acid sequence
10 EGGDVDDLLDMI.

4. A screening method according to any of claims 1 to 3 wherein said binding partner or said test agent is a polypeptide having a amino acid sequence which shares at 15 least 50% homology with DNA-PK_{cs}.

5. A method according to any of claims 1 to 4 wherein said binding partner or said agent is in a test sample.

20 6. A method according to claim 5 including quantifying the amount of the agent or binding partner in the sample.

7. A screening or assay method according to any one of the preceding claims including purifying and/or isolating 25 a test agent or a test binding partner from a mixture or extract.

8. A screening or assay method according to any one of the

preceding claims including determining the ability of one or more fractions of a test mixture or extract to bind to said peptide.

- 5 9. A method according to any one of the preceding claims comprising labelling one of said test agent or said binding partner and said peptide with a detectable label, immobilising the other on a solid support and bringing the test binding partner or test agent and the peptide 10 into contact.
10. A peptide consisting essentially of an amino acid sequence which shares at least 70% homology with an amino acid sequence of a Homology Region of Ku80 or Ku70 as 15 shown in Figure 1.
11. A peptide according to claim 10 consisting essentially of an amino acid sequence of a Homology Region of Ku80 or Ku70 as shown in Figure 1. 20
12. A peptide according to claim 10 consisting essentially of an amino acid sequence which differs from a Homology Region of Ku80 or Ku70 as shown in Figure 1 by deletion of one or more amino acid residues. 25
13. A peptide consisting essentially of the amino acid sequence EGGDVDDLLDMI or a sequence variant which shares at least 50% homology thereto, said peptide having the

ability to bind DNA-PK_{cs}.

14. A peptide according to claim 13 consisting essentially of a variant of the amino acid sequence EGGDVDDLLDMI as shown in Figure 6.

5 15. A peptide according to any one of claims 10-14 including one or more heterologous amino acids joined to the specified peptide.

10

16. A polypeptide comprising a peptide of any one of the preceding claims, said polypeptide having 394 amino acids or less.

15 17. A polypeptide according to claim 16 selected from the group consisting of Ku80ΔN1, Ku80ΔN2, Ku80ΔN4, Ku80ΔN5, Ku80ΔN6, Ku80ΔN7 and Ku80ΔN8 as shown in Figure 2, said polypeptide having the ability to bind to DNA-PK_{cs}.

20 18. An isolated nucleic acid encoding a peptide or polypeptide in accordance with any one of claims 10 to 17.

19. A vector comprising nucleic acid according to claim 18.

25

20. A vector according to claim 19 comprising one or more control sequences operably linked to said nucleic acid.

21. A host cell comprising a vector according to claim 19 or
claim 20.

22. A transgenic animal comprising a host cell according to
5 claim 21.

23. An agent or binding partner identified by a method
according to any one of claims 1 to 9.

10 24. An agent or binding partner according to claim 23
comprising either;

- i) a peptide consisting of the amino acid sequence
EGGDVDDLLDMI, or
- ii) a peptide fragment of DNA-PK_{cs}.

15

25. Use of an peptide according to any one of claims 10 to 17
for determining the presence in a test sample of an agent
or binding partner which has the ability to bind to Ku70
or Ku80.

20

26. Use of an agent which binds a peptide according to any
one of claims 10 to 17 for determining the presence in a
sample of a Ku70 or Ku80.

25 27. A method of purifying a homologue, orthologue or variant
of DNA-PK_{cs} having at least 50% sequence homology to human
DNA-PK_{cs},
said method including contacting the homologue,

orthologue or variant with a peptide according to claim 13 or claim 14.

28. A method of purifying a polypeptide comprising the
5 sequence EGGDVDDLLDMI,
the method including contacting material containing the polypeptide with DNA-PK_{cs}.

29. A method of producing a peptide or polypeptide according
10 to any one of claims 10 to 17 comprising transforming a vector according to claim 19 or claim 20 into a host cell, culturing the host cell, and recovering the peptide or polypeptide from the host cells or the surrounding medium.

15

30. An antibody directed to a peptide according to any one of claims 10 to 17.

31. An antibody directed to an agent or binding partner
20 according to claim 23.

32. A method of purifying and/or isolating a polypeptide according to any one of claims 10 to 17, comprising contacting said polypeptide with an antibody according to
25 claim 30.

33. Use of a peptide according to any one of claims 10 to 17 in a method of designing a peptide or non-peptidyl

mimetic of said peptide.

34. Use of an agent according to claim 23 in a method of obtaining a peptidyl or non-peptidyl mimetic.

5

35. A method of designing a mimetic of a peptide according to any one of claims 10 to 17 which modulates the interaction of Ku70 or Ku80 with a binding partner, the method comprising:

10 (i) analysing said peptide to determine the amino acid residues essential and important for the activity to define a pharmacophore; and,

(ii) modelling the pharmacophore to design and/or screen candidate mimetics having the ability to modulate the interaction of Ku70 or Ku80 with a binding partner.

15

36. A mimetic obtained by a method of claim 35.

37. A peptide according to any one of claims 10 to 17 for use
20 in therapy.

38. An agent or binding partner according to claim 23 for use in therapy.

25 39. A pharmaceutical composition comprising a peptide according to claim 37, an agent according to claim 38 or an antibody according to claim 30 and a pharmaceutically acceptable carrier.

40. A method of treatment of a condition associated with a defect or disorder in DNA repair or cell cycle control comprising administration of a peptide according to claims 10 to 17, an agent or binding partner according to 5 claim 23, an antibody according to claim 30 or a composition of claim 39.

41. Use of a peptide according to claims 10 to 17 or an agent or binding partner according to claim 23, or an antibody according to claim 30 in the manufacture of a medicament 10 for use in the treatment of a defect or disorder in DNA repair or cell cycle control.

hsku70	1	...MSGWESYYKTEG...DEEAEEEQE
ggku70	1	MEMWVLGEVGMAVLSAAMADWVSYRGDGPDEED DGEQQEE
raku70	1	MDQPWMRQDD E SDDE
dmku70	1	MSTWINPENDV DLLSGSE
scku70	1	MRPVVTNAFGNSGELN
hsku80	1	...
ceku80	1	...
scku80	1	...
RE		
hsku70	22	ENLEAS G DY K YS GRD SL I FLV D ASKAMFES.Q SE DEL T PFDM
ggku70	43	EGPEAV A DY R FS GRD SL I FLV D ASKAMFEPYENE E AA T PFDM
raku70	17	SSTVDF G Q...AVDG T FL I DATEGM E EVD G D...TA F MQ
dmku70	18	DEEDVSMKRD Y H CRE A T LF V V D ANL Q ...
scku70	16	D Q V D E K GYR K FD I HE G IL C IE L SET M FE K ESS D LD Y K S PL E
hsku80	1	MVRSGNKAA V V L CM D V G FT M S.N.SIP G IES P FE Q
ceku80	1	MP P P K KVSPGI T V I I D A GSNM S VK.D T E T G KS A FE N
scku80	1	MSSE S TT F IV D V S PSMM K N.NNVSKSMAY LE

2/11

hsku70	63	SITQCISSDRTDLIAVVFYGT
ggku70	85	TMOCIRNVTSSDKDLSSVVFYGM
raku70	52	CTKAAKSMTMLNKTSSPKDLVGTITFGT
dmku70	53	AINTTRIAFISGILVNDKDLTGLIFANTKHSPPPLEASA
scku70	58	TESTIDEMSQLVTRPGTATGCYFY
hsku80	34	AKKVTITMENVQRQVFAENKDETALVILEGT
celku80	36	ATNAADWIVSRKDPPEIFESVMay
scku80	31	YTLINKSKKSRKTDWITSCYLANCPVSENSQEIIPN.VF

hsku70	91	... EKDKN SVNEKNT YVVIQE
ggku70	113	... ENNKSADFKHTYVVLQE
raku70	80	... DKKDNPNREKNVYVVLQD
dmku70	92	... LDNIVMPDNCAVFLPLRQL
scku70	84	... YCNREDAKEGTYELFP
hsku80	75	... QNITVHRHMLPD. FDILLED
celku80	78	... NGVQCQNEKFTPASEFDHL
scku80	67	... QIQSFLAPVTTATIGEIKR

Figure 1 (continued)

hsku70	130	RFQDMMGGSDYSLSEVILWVCANLFS
ggku70	152	LFRETFGHNADYSLGEALWACSNLF
raku70	118	KEKQEQYGHG.NVNADVILWTCALM
dmku70	132	VY.GLAEPDGRGRFDLMI
scku70	125	RLCTEMLEKC.LYDYFMFQQTGSEKQVR
hsku80	114	SVLIFTFMLDTE.SMDVI.QH.ETIGK
ceku80	120	KEKRIEIFTDLSS.RFSKSQLD
scku80	109	DLII.AVAVLKDDQI.ESY

hsku70	157	AKASRARTKAG.DLRDT
ggku70	179	MLFTNEDNPHGNDSAKAKLARTRAG.DLRDT
raku70	143	DLIIQSSAKAKLARTRAG.DLRDT
dmku70	159	DLIIQSSAKAKLARTRAG.DLRDT
scku70	158	MLFTDIDKPKQEAQDIDERARIRRLTIDFDN
hsku80	151	PGQKQLSNKRVFLFTHSLKKCDTSLQFF
ceku80	159	FLPPFSLGKEDCGSGDRGDGPFRILGG.HGP
scku80	151	MIIGIDENPEYPAASRVAAELAETLEG.RTY

Figure 1 (continued)

hsku70	197	G IFLDL M HKKPGG . . .
ggku70	219	G ITLDL M HKKPGG . . .
raku70	182	G IELDL V HLKPPGDKKERPQILYKNTIVTD . . .
dmku70	198	EEEFHVTPM . . .
scku70	200	KVNEATE . FIGYADKPEDNE .
hsku80	191	SFP . . .
ceku80	198	TFQNVAKMLSTFQARQKSERKYNK . .
scku80	187	R TYNMELLVEITSPATSIVVKPVR . .

hsku70	232	HF E ESSKLED . . .
ggku70	254	QPDES G KLEH . . .
raku70	219	P EASDK M EE . . .
dmku70	235	DAQMLRE . . .
scku70	240	FDGPs T KPIDAKYIKSRILRKKEVKRIMFQCP L ILDEKTNFI
hsku80	228	SESLRK L CVFKKIERHS . . .
ceku80	238	KSEK S TALLKFKNADSEGNNEMVRLEQMHVETDEVAP K DEFKS . . .
scku80	227	T S NP S GSM . . .

5/11

nsku70	271	VGTYNL VQKALKPPPIKLT
ggku70	293	VGVYNL EOKAYKPYPVKL
raku70	257	VSVSYNL VIRPTGKPATTRL
dmku70	272	VOYYNYFQRAYPRKVQI
scku70	282	VGVGYTMYTHEKAGVRYKLVYEHEDIRQEAYSKR
hsku80	267	TIHQERVKKTWTVVDAKTIKEEDIQKETVYCLNDDETE
celku80	280	PVLEEKPK
scku80	262	VEVEDSQKKER YVGVKSIIEYEIHNEGNGNKKNVSEDQSGSSY

hsku70	308	GG	...	LLPSDTKRSQIYGRQITILEKEETEELKR	...	LLPSDTKRAQTYGNRQITAMEKEETEVKR	...	LLPSDTKRAQTYGNRQITAMEKEETEVKR	...
ggku70	330	GS
raku70	294	AE
dmku70	309	KDDGSQDTHEEYQTIKVTGGWYTCNVGER
sclku70	324	GED
hsku80	305
celku80	309
scku80	304	IPV

hsku70	340	E	DDPGLMLMGFKPLV.LIKKHHYLRPSL
ggku70	362	E	DSPGLFLIGFKPLS.MLKQHHHTRPSQFM
raku70	326	M	APPGLQLLGFKPLS.YLEKQPHVRPSHEVYPD
dmku70	351	L	HKPQMMLLGFKHRS.SLPEVSYIKPANFMYPD
scku70	361	.	AFLKITTIGFRSSSKSIHYENNIDKSSEFIVPD
hsku80	332	K	YKSEGKCESVVLGFCKSSQVQRRFFMGNQVLFKAARD
ceku80	339		DHNNEGQTTGGVLKLIQFTKRANTLDSYLLDASAKTVLPALN
scku80	337	E	SFPGLDILRGFLNR.EALPRYFLTSESSSEITADTR
hsku70	373	S	SLVIGSSITLFSALLIKCLEKEVAALCRYTPRNITPPYFVALV
ggku70	395	S	SLVTGSTTLFNALLMKCLEKEVIMALCRYTIAARRNTPPRIVALI
raku70	359	G	GSVRGSSTRLFAALLQSCLRRVAPICFWISRAAQAPKLVYLL
dmku70	384	Q	QSTIGSKRILFERCLVRDKIAICLFMCKRKSIIPRYVALV
scku70	392	A	AKYEGSIRTLASLLKILRKKDKKIAILWGKLKSNSHPSLYTLS
hsku80	370	.	DEAAAVALSSLTHALDDLMVATVRYAYDKRANPQVGVAF
ceku80	381	S	SPKSGATKATVSLSLRVAATCRYTFHAKSHVQQLIAALL
scku80	371	L	LGCQSDLMAFSALVDVMLEVARYVSKKDSEVNMCALC

			HR4				
hsku70	415	PQEELDDQKIQ	VTPPGFQLVFLPFADDKRKMPF				
ggku70	437	PQEEEVIDEQKVQ	TAPPGFHITFLPYADDKRNVDF				
raku70	401	AQEERERDPHGLQ	MVPPGFHVQLPFSDDRRRLQQA	I			
dmlku70	426	PVEAPDNGEDKNYRSLLCGDGFKIV	VYLPEAKHIRHLDLQDW				
scku70	434	PSSVKDYNE	GFYLYRPPFLDEIRKFP	SIL			
hsku80	410	PHIKHNYEC	YVOLPFPMEDILRQYMFSS				
ceku80	423	PHQDEETGVFY	LSVKLPPFSDDMRTLKEPKFS				
scku80	413	PVLTIEHSNINSEKKFVKS	LTCRLLPEAEDERVTDFPKLL				
			HR5				
hsku70	449	TEKIMATPEQVGKMKAI	VEKLRFTY				
ggku70	471	TEKVPANREQVDKMKGIIQ	KLRFKY				
raku70	436	QEGTTKATPGLVALAREMAE	KLRFTY				
dmlku70	467	NNTENTADEQKVEFFQKIIK	KLRVDY				
scku70	463	SYDDGSEHKLDYDNMKKKV	NLRDGY				
hsku80	439	TSIIMGYF					
ceku80	455	KNSKKYAPTEAQLINAVDALITDSMSLAKKDEKTD	LEDLE				
scku80	452	QVDEDDILNPKIVAQQLSAVDDLIDCMQLQ	DEI	S			
		QFVDSM					

hsku70	474	RSDSFENPVLQ	QHF	RNLEALALDIMEPE	QAVDLT	EPKVEAMN
ggku70	496	RTDSFENPVLQ	QHF	RNLEALALDML	EPKTEEMS	
raku70	462	HPDKFENPELQG	FWSC	LEALALDRDDAE	HPKDYTRP	DHEKMK
dmku70	493	QPNLINDPSL	QANLIALS	LD	STD	TKGLD
scku70	495	NPSDEKNP	YK	YKVLHDYLLQ	ETT	EDENE
hsku80	478	PTTKIPNPRF	QRLF	QCLLHRA	LHPR	EPPLP
ceku80	491	VEGGMSDPKL	LQM	QCHFLKSL	VLH	PN
scku80	482	DTDELPE	PI	IGNYYQPIGEV	TTLD	PSL

hsku70	516	KRLGSLVDE	FEKELVY	P	PDYN	PEGKVT	.KR			
ggku70	538	RRI	GNL	V	PDYS	PEGKAA	.KR			
raku70	504	AKAGE	MDA	FEAAFP	DGCSATT	AGSR	KR			
dmku70	534	KRIEKLI	PDYE	.MFAP	EAEP	PKKRAAKST				
scku70	533	DRMMREDDSI	RKLY	YI	RNKI	LES	EKSEDPTIQRILNK			
hsku80	516	LNPPAEV	TKS	Q	TEAKKKDQV	T	TAQE	IFQ		
ceku80	533	MAPKRRV	AEHPEI	FQKLGREFNLQ	QKTKRER	W	VEPE	DL		
scku80	524	LRI	P	PTMVYRQQ	V	LEWIHQL	.M	INDS	REFE	PELPDSL

Figure 1 (continued)

hsku70	544	KHDNEGSGSKRPKVEY . . . SEEELKTHISKGTLGKFTVPMILK .
ggku70	566	KQAGDAAQAEKRPKIEI . . . SEDSRSYVQNGTILGKLTVSAIK .
raku70	533	TQAGEGGQAKKARSSENQGSNVDIREEAKRGKLIASLTVSVLR .
dmku70	562	TAGASGPKMAKINDDQLKEFEVKSILNKDEALTSCCTAAQLH .
scku70	569	YVKIWNMFYKKFENDDNISTKEEKPKFDKKPKFNI . . .
hsku80	556	DNHEDGPTAKKLKTEQGGAHFSVSSIAEGSVTSGSV . . . N
ceku80	575	QTMISENTEKKQNMQTQPDDEVDDGASQKKKKPNAKKI . . . T
scku80	563	KNKISPYTHKKFDSTKLVEVLGIKKVDKLKLDDSELKTELERE .

hsku70	583	EACRAYGLKSGLKKQELLEAILTGHKFQD . . .
ggku70	605	DTCRYHYGLRSGGKKQELIDALTEYFSGR . . .
raku70	574	DFCKQEGLRCPSKKAELVDCIHKHLKL . . .
dmku70	604	ILQHHFDVTMPKSSKKAKLVAKIEELHK . . .
scku70		
hsku80	594	PAENFRVILVKQKIKASFEEASN . . . QLINHIEQFL . DTNETPYF
ceku80	613	RKEEVQMDIMEDGASRVCVCSKIEEMITSNTCKFQPNGAVTEFF
scku80	605	KIPDIEETILKRGEOQHSRGSPNN SNN . . .

Figure 1 (continued)

hsku70	675	WEIVVQDGITLITKEEAKKEFLAPKDKPSGDTA
ggku70	696	VLSEEKSCNPIDSS...
raku70		VSMSEVSVANAA
dmku70		
scku70		
hsku80		
ceku80		
scku80		

SUBSTITUTE SHEET (RULE 26)

Figure 1 (continued)

hsku70
ggku70
raku70
dmku70
scku70
hsku80 717 AVFEEGGDVDDLLDMI
ceku80 723 EFWEED
scku80

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 00/01245

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G01N33/68 C07K7/08 C07K14/47 A01K67/027 C07K16/18
A61K38/10 C12N9/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, BIOSIS, CHEM ABS Data, STRAND

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category ^o	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 30903 A (CANCER RES CAMPAIGN TECH ;DOWNS JESSICA ANNE (GB); JACKSON STEPHEN) 16 July 1998 (1998-07-16) claims 11-14 ---	1-12, 15-41
X	WU X ET AL: "PROTEIN-PROTEIN AND PROTEIN-DNA INTERACTION REGIONS WITHIN THE DNA END-BINDING PROTEIN KU70-KU86" MOLECULAR AND CELLULAR BIOLOGY, US, AMERICAN SOCIETY FOR MICROBIOLOGY, WASHINGTON, vol. 16, no. 9, September 1996 (1996-09), pages 5186-5193-93, XP000877319 ISSN: 0270-7306 abstract --- -/-	1-12, 15-41

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- A• document defining the general state of the art which is not considered to be of particular relevance
- E• earlier document but published on or after the international filing date
- L• document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- O• document referring to an oral disclosure, use, exhibition or other means
- P• document published prior to the international filing date but later than the priority date claimed

- T• later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- X• document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- Y• document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- &• document member of the same patent family

Date of the actual completion of the international search

20 September 2000

Date of mailing of the international search report

11/10/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Hart-Davis, J

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 00/01245

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOIKE MANABU; MIYASAKA TAKASHI; MIMORI TSUNEYOSHI; SHIOMI TADAHIRO: "Subcellular localization and protein-protein interaction regions of Ku proteins" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 252, 27 November 1998 (1998-11-27), pages 679-685, XP002143822 abstract --- CARY ROBERT B; CHEN FANQING; SHEN ZHIYUAN; CHEN DAVID J: "A central region of Ku80 mediates interaction with Ku70 in vivo" NUCLEIC ACIDS RESEARCH, vol. 26, 15 February 1998 (1998-02-15), pages 974-979, XP002143823 abstract --- GELL DAVID; JACKSON STEPHEN P: "Mapping of protein-protein interactions within the DNA-dependent protein kinase complex" NUCLEIC ACIDS RESEARCH, vol. 27, 1 September 1999 (1999-09-01), pages 3494-3502, XP002143824 the whole document --- SINGLETON B K; TORRES-ARZAYUS M I; ROTTINGHAUS S T; TACCIOLI G E; JEGGO P A: "The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit" MOLECULAR AND CELLULAR BIOLOGY, vol. 19, May 1999 (1999-05), pages 3267-3277, XP000929550 the whole document --- WEDRYCHOWSKI, ANDRZEJ ET AL: "Identification of proteins binding to interferon-inducible transcriptional enhancers in hematopoietic cells" J. BIOL. CHEM. (1992), 267(7), 4533-40, XP002143825 page 4535, column 2, paragraph 2; figure 3 --- MYUNG, KYUNGJAE ET AL: "KARP-1: a novel leucine zipper protein expressed from the Ku86 autoantigen locus is implicated in the control of DNA-dependent protein kinase activity" EMBO J. (1997), 16(11), 3172-3184, XP002143826 page 3176; figure 5 ---	1-12, 15-41 1-12, 15-41 1-41 1-12, 15-41 13-24, 29-34, 36-39 13-24, 29-34, 36-39
		-/-

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 00/01245

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MIMORI, TSUNEYOSHI ET AL: "Isolation and characterization of cDNA encoding the 80-kDa subunit protein of the human autoantigen Ku (p70/p80) recognized by autoantibodies from patients with scleroderma-polymyositis overlap syndrome" PROC. NATL. ACAD. SCI. U. S. A. (1990), 87(5), 1777-81 , XP002143827 page 1779; figure 3 ---	13-24, 29-34, 36-39
X	WASSERMAN ROBERT; GALILI NAOMI; ITO YOSHINORI; REICHARD BETTY ANNE; SHANE SARA; ROVERA GIOVANNI: "Predominance of fetal type DJ-H joining in young children with B precursor lymphoblastic leukemia as evidence for an in utero transforming event" JOURNAL OF EXPERIMENTAL MEDICINE, vol. 176, 1992, pages 1577-1581, XP000933767 the whole document ---	13-24, 29-34, 36-39
X	US 5 635 597 A (BARRETT RONALD W ET AL) 3 June 1997 (1997-06-03) table 1 ---	13-24, 29-34, 36-39
X	DATABASE EMEST_ROD3 EBI, Hinxton, UK; Accession Number AI181344, 13 October 1998 (1998-10-13) MARRA M ET AL.: "Mus musculus cDNA clone IMAGE:1395081 5' similar to TR:Q62084 Q62084 PHOSPHOLIPASE C NEIGHBORING " XP002143829 the whole document -----	13-24, 29-34, 36-39

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 23,31,34,36,38,40-41 (in part)

Present claims 23,31,34,36,38, as well as 40 and 41 in so much as they depend on claim 23, relate to a compound defined by reference to a desirable property, namely its ability to bind to a peptide derived from polypeptides of the Ku70 or Ku80 families, binding being determined by an in vitro (claims 1-9) or modelling (claim 35) screening method.

No technical features of the compounds are present in the above-mentioned claims which would lead to this desirable property, the technical features formulated so as to permit the execution of a meaningful search. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found for only a very small proportion of the substances which could fall within the scope of these claims. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. No means are present in the above-mentioned claims by which compounds known in the prior art could be distinguished from novel substances. No definition of the subject matter for which protection is sought is therefore derivable from these claims (Article 6 PCT) or the description (Article 5 PCT). Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search for claims 23,31,34,36,38, 40 and 41 has been restricted to the substances which are clearly defined and supported by the description, namely polypeptides of the Ku70 or Ku80 families and peptides derived therefrom such as EGGDVDDLLDMI.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 00/01245

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9830903	A 16-07-1998	AU 5568198 A		03-08-1998
		AU 5568498 A		03-08-1998
		EP 0966683 A		29-12-1999
		EP 0965040 A		22-12-1999
		WO 9830902 A		16-07-1998
		GB 2321702 A, B		05-08-1998
		GB 2322193 A, B		19-08-1998
		GB 2329469 A, B		24-03-1999
		GB 2329248 A, B		17-03-1999
US 5635597	A 03-06-1997	NONE		

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.